Collaborative knowledge creation: Evidence from Japanese patent data

Mori, Tomoya and Sakaguchi, Shosei

Institute of Economic Research, Kyoto University, Department of Economics, University College London

29 August 2018

Collaborative Knowledge Creation: Evidence from Japanese Patent Data[*]

Tomoya Mori[†,‡] Shosei Sakaguchi[§]

February 21, 2019

Abstract

This paper presents micro-econometric evidence for collaborative knowledge creation at the level of individual researchers. The key determinant for developing new ideas is the exchange of differentiated knowledge among collaborators. To stay creative, inventors seek opportunities to shift their technological expertise to unexplored niches by utilizing the differentiated knowledge of new collaborators. Moreover, the knowledge stock of an inventor, proxied by the scope of an inventor's past research, has positive and negative effects on their productivity. This is because it facilitates successful collaboration; however, simultaneously, the dependence on older knowledge hinders invention possibly due to the obsolescence and exhaustion of niches by imitation.

Keywords: Knowledge creation, Collaboration, Differentiated knowledge, Patents, Technological novelty, Technological shift, Recombination, Network

JEL Classification: D83, D85, O31, R11, C33, C36

---

[*]Earlier versions of this paper were presented at Jinan University, Keio University, Kyoto University, RIETI, the University of Tokyo, the 32nd ARSC meeting at Nansan University, the 13th UEA meeting at Columbia University, and the 2nd CURE meeting at Singapore Management University. The authors are particularly indebted to Marcus Berliant, Jorge De la Roca, Gilles Duranton, Masahisa Fujita, Tadao Hoshino, Wen-Tai Hsu, Daishoku Kanehara, Shin Kanaya, Kentaro Nakajima, Ryo Nakajima, Yoshihiko Nishiyama, Ryo Okui, and Akihisa Shibata for their constructive comments. This research was conducted as part of the project "An empirical framework for studying spatial patterns and causal relationships of economic agglomeration" undertaken at RIETI. It received partial financial support from JSPS Grants-in-Aid for Scientific Research Grant Numbers 15H03344, 16K13360 and 17H00987.

[†]Institute of Economic Research, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan. E-mail: mori@kier.kyoto-u.ac.jp.

[‡]Research Institute of Economy, Trade and Industry (RIETI), 11th floor, Annex, Ministry of Economy, Trade and Industry (METI) 1-3-1, Kasumigaseki Chiyoda-ku, Tokyo 100-8901, Japan.

[§]Department of Economics, University College London, Gower Street, London WC1E 6BT, UK. E-mail: s.sakaguchi@ucl.ac.uk.

1 Introduction

Knowledge creation has been a key factor in various aspects of economic modeling. Some new ideas result in innovations and fuel economic growth.[1] The structure of market and that of competition may depend on the extent of diffusion and imitations of technologies.[2] The concentration of R&D activities is the defining feature of the largest cities.[3] However, knowledge creation at the ultimate micro level of individual inventors has been abstracted in these strands of the literature. The corresponding empirical studies are necessarily scarce. This study investigates data on Japanese patents applied between 1995 and 2009. It focuses on causalities in collaborative knowledge creation primarily based on the microeconomic model proposed by Berliant and Fujita (2008) which, to our knowledge, is the first attempt to formalize active knowledge creation by individual inventors.[4,5]

The productivity of inventors is volatile. In our data, we find substantial downward pressure on inventor productivity as well as substantial churning of the productivity ranking of inventors over time.[6] Specifically, less than half of inventors with above-median productivity in a given period maintain at least the same relative productivity in the next period. In this case, some top inventors stay highly productive, whereas some inferior inventors overthrow superior ones and climb the productivity ladder.[7]

The extant literature provides an explanation for the declining trend of inventor productivity. Inventors have an incentive to exploit their expertise on the existent technologies through learning-by-doing (Horii, 2012). However, once made public, technologies face incessant innovations by which new technologies replace old ones (e.g., Grossman and Helpman, 1991; Klette and Kortum, 2004; Lucas and Moll, 2014). Moreover, publicized technologies attract imitations that deprive opportunities to profit by refining them (e.g., Chu, 2009; Cozzi and Galli, 2014). The latter negative

[1]See, for example, Shell (1966); Romer (1990); Grossman and Helpman (1991); Aghion and Howitt (1992); Kortum (1997); Klette and Kortum (2004); their extensions (e.g., Lentz and Mortensen, 2008; Akcigit and Kerr, 2018; Cai and Li, 2019), and the studies that have focused on knowledge diffusion (e.g., Scherer, 1982; Jovanovic and Rob, 1989; Coe and Helpman, 1995; Lucas and Moll, 2014; Perla and Tonetti, 2018).

[2]See, for example, König et al. (2019); Panebianco et al. (2016) on the interdependencies among R&D collaboration, technology diffusion, and product market competition among firms; related studies in the context of economic growth (e.g., Yang and Maskus, 2001; Glass and Saggi, 2002; Tanaka et al., 2007); and those in conjunction with the role of patent system (e.g., Grossman and Shapiro, 1978; Chang, 1995; Matutes et al., 1996; Schotchmer, 1996).

[3]For example, Duranton and Puga (2001); Bettencourt et al. (2007); Davis and Dingel (2019).

[4]We focus on knowledge creation and not innovation as patents may not induce innovation.

[5]In an alternative model by Weitzman (1998), a new idea induces the development of other new ideas if recombined with the existing ideas. In the model by Olsson (2000, 2005), a new idea is created from a convex combination of existing ideas or existing and impromptu ideas. Unlike the model of Berliant and Fujita (2008), however, knowledge creation in these models is passive and accidentally.

[6]Multiple measures of inventor productivity are considered. See sections 2.1 and 5.2 for the details.

[7]Here, three five-year periods between 1995 and 2009 are considered. See Section 2 for the details.

effects eventually dominate because learning-by-doing is subject to decreasing returns (Horii, 2012).

How can inventors achieve high productivity in these circumstances? Horii (2012) proposed a model of innovation associated with technological shifts. In his model, consumers wish to satisfy an infinite range of wants. This induces an inventor to seek an unexplored technological niche wherein they can create demand for new products realized by new technology. While this model lacks a micro mechanism for the technological shifts, it is complemented by the model of Berliant and Fujita (2008).

In the Berliant–Fujita (BF) model, agents communicate via common knowledge and invent in pairs by utilizing their mutual differentiated knowledge. Here, an appropriate balance between common and differentiated knowledge facilitates collaborative innovation. A longer duration of collaboration by the same pair increases their common knowledge while decreasing their mutual differentiated knowledge. Simultaneously, this accumulates differentiated knowledge between them and the remaining agents. Agents optimally choose the set of their collaborators and time allocation for each collaboration to achieve the best knowledge composition and maximize their total output.

Given these facts and theoretical backgrounds, three regression models are developed. The first model is a simplified representation of the BF model. It expresses knowledge creation by an inventor and their average collaborator, where the inventor's optimal choice of polyadic collaboration is implicit. In this model, we focus on the differentiated knowledge of collaborators while abstracting from the common knowledge and the differentiated knowledge of the inventor. This is because it is the key source of new ideas in the theoretical model. It is quantified in terms of the collaborator's output, *excluding* the patents jointly developed with the inventor. Our baseline results indicate that a 10% increase in collaborators' differentiated knowledge for an inventor raises their research output by 1.7–3.4%. This implies positive but decreasing returns of this knowledge and is consistent with the results of the theoretical model.

The second model decomposes the contribution by collaborators' differentiated knowledge to the research output of an inventor (computed from the regression of the first model) into the fractions accruing to the quality and quantity of their output. We find that the contribution is mostly dedicated to increasing the quantity rather than the quality of output if the patent quality is measured by the cited count. However, approximately 65% of the contribution accounts for increasing the quality of output if the patent quality is measured by technological novelty. Accordingly, a major role of collaboration is to induce technological shifts of an inventor to a new niche.

The third model investigates factors that determine the amount of differentiated knowledge that each inventor obtains from their collaborators. We find that a more active recombination has a positive selection effect on collaborations, leading to a set of new collaborators with a larger average differentiated knowledge. Our baseline results indicate that a 10% increase in the number of new collaborators of an inventor raises the average differentiated knowledge of collaborators by 12–17%. This eventually raises the inventor's research output by 3–8%. Furthermore, a larger research experience, measured by the scope of an inventor's past research and broadly interpreted as knowledge stock, exhibits a positive effect on attracting collaborators with more differentiated knowledge even though it has a negative direct effect on inventor productivity as discussed above.

In these regressions, we control for individual fixed effects by exploiting panel data and a variety of establishment, firm, industrial, and other local factors. However, we face identification problems due to network endogeneity that arises from endogenous collaborations of inventors. The identification and estimation of models with endogenous networks are substantial challenges in the literature on network econometrics (e.g., Jackson et al., 2017). The most common way to tackle network endogeneity in the context of our problem is to consider a network formation model to identify and estimate the parameters of interest.[8] However, because the BF model provides no simple econometric model of network formation as will be clear in Section 3, this approach fails in our case. Therefore, we propose an alternative approach to manage endogenous regressors for an inventor through instrumental variables constructed from information on their indirect collaborators.

Typically, using more distant indirect collaborators to construct the instrumental variables is double-edged. This is because it not only reduces the reflection problem (see, e.g., Bramoullé et al., 2009) but also makes the instrument weaker. However, we benefit from a special situation in which the relevance of instruments is extrinsic to the inventor network as it comes from the positive assortative matching by productivity among firms and workers.[9] The matching is essentially exogenous to individual inventors given that it occurs prior to the formation of a research network and is based on more diverse aspects of firms' profit maximization than on R&D activities. Consequently, the relevance of instruments is maintained even when the information of only distant indirect collaborators is used without resorting to external variations (as in, e.g., Azoulay et al., 2010; Waldinger, 2010, 2012). This is possible as long as the assortative matching simultaneously affects the indirect collaborators and the targeted inventors.

In the extant literature, the most closely related study is by Akcigit et al. (2018). They

[8]See, for example, Goldsmith-Pinkham and Imbens (2013); Hsieh and Lee (2016); Comola and Prina (2017); Li and Zhao (2016); Patacchini et al. (2017); Johnsson and Moon (2019). Another typical approach assumes an exogenous network (e.g., Bramoullé et al., 2009; Akcigit et al., 2018).

[9]See, for example, Mori and Turrini (2005); Mendes et al. (2010); Bartolucci and Devicienti (2013); Behrens et al. (2014); Eeckhout and Kircher (2018); Gaubert (2018). In Section 6.3, we add supportive evidence from the financial and ownership data of firms in Japan.

estimated a reduced-form model of team-level innovation similar to the knowledge creation model of Berliant and Fujita (2008). Their key explanatory variables are the quantity and quality of interactions within a team. The crucial differences from our approach are that their "interaction" effect results from the past experience of the team leader and not from their current collaborators (as in our case). Furthermore, the collaboration network in their study is assumed to be exogenous.

Other related studies include those on positive externality in knowledge creation (e.g., Azoulay et al., 2010; Waldinger, 2010, 2012; Iaria et al., 2018) and on knowledge diffusion (e.g., Jaffe et al., 1993; Thompson and Fox-Kean, 2005; Murata et al., 2014; Kerr and Kominers, 2015). Our paper is closer to the former, which tests if superior researchers positively impact their collaborators' productivity. A crucial difference is that we are explicit about the causal channel through the exchange of differentiated knowledge between collaborators, whereas the mechanisms are abstracted in the form of spillover effects in their studies. The latter studies concern the distance and routes through which knowledge spreads and not how they are created.[10]

The rest of the paper is organized as follows. Section 2 presents key observations on the dynamics of knowledge creation and inventor productivities. Sections 3 and 4 describe the BF model and the corresponding regression models, respectively. Section 5 presents the data, Section 6 discusses the identification strategy, and Section 7 presents the baseline regression results. Section 8 provides a series of robustness checks, and Section 9 presents the conclusion.

2 Facts

We make three observations on patent development in Japan to guide our analyses.

2.1 Productivity of an inventor

Our panel data comprise three periods, each of which aggregates five consecutive years: periods 0, 1, and 2 include years from 1995 to 1999, 2000 to 2004, and 2005 to 2009, respectively. We focus on the balanced set I of 107,724 inventors, each of whom participate in at least one patent in each period.

Let $\mathcal{G}_{it}$ be the set of patents in which inventor i participates in period t, and G_j for $j \in \mathcal{G}_{it}$ is the set of inventors who participate in patent j. Denoting the output of patent project j by a scalar $g_j > 0$, *the productivity of the inventor* i is defined in terms of the total output of patents in which they participated, with the output of each patent being

[10] See Breschi et al. (2003); Garcia-Vega (2006); Østergaard et al. (2011); Huo and Motohashi (2015); Inoue et al. (2015) for quantifications of common and differentiated knowledge in developing new ideas.

discounted by the number of inventors involved in the patent:

$$\bar{y}_{it} = \sum_{j \in \mathcal{G}_{it}} g_j / |G_j| \tag{2.1}$$

where $|G_j|$ means the cardinality of set G_j. (Hereafter, the expression $|X|$ for any set X means the cardinality of X.)

We consider two quality measures of inventor productivity in our baseline analysis (and three others to check the robustness of results in Section 8.3). One is the cited count, where g_j represents the count of citations that patent j received in three years of publication. The other is *novelty*, where g_j represents the *degree of technological novelty* of patent j as defined by the reciprocal, $1/r_j$, of the order, $r_j = 1, 2, \ldots$, of j in terms of its application date among all the patents classified in the same technological category as j.[11,12] The technological category of a patent is identified by the "subgroup" of the International Patent Classification (IPC) in the baseline analyses.[13]

2.2 Dynamics of the relative productivities of inventors

Let $I_t^{\mathrm{TOP}}(x)$ represent the set of inventors in the top x% in I in terms of their productivity in each period $t = 0, 1$, and 2. The set of inventors in each 5% interval of the productivity percentiles from 0% to 100% can then be expressed by $\Gamma_t(x) \equiv I_t^{\mathrm{TOP}}(x) \backslash I_t^{\mathrm{TOP}}(x-5)$ for $x = 5, 10, \ldots, 100$, where "\" is a set difference operator. Call $\Gamma_t(x)$ the (*productivity*) *class* x of inventors in period t.

For classes $x = 5, 10, \ldots, 100$ under citation and novelty-adjusted productivities, the height of each blue bar in Panels (a) and (b) in Figure 2.1, respectively, indicates the share of inventors of class x in period 0 who stay at least in the same class x' ($\leq x$) in period 1. The graphs reveal a clear pattern:[14]

Observation 1 (Churning of relative productivities) *Under either measure of productivity, less than half of inventors above the median productivity $x < 50$ in period $t-1$ remain at least as productive in period $t \in \{1, 2\}$. This indicates a strong pressure to prevent inventors from maintaining their relative productivity. In other words, a sizable proportion of inferior inventors replaces superior ones in their productivity ranking in each period.*

[11] Our data include all the patents applied in 1993 and thereafter as well as some older applications published in 1993 or later. By construction, our measure of novelty overstates the novelty in technological categories defined before 1993. However, because our regression analyses use novelty data from 2000 and later (i.e., periods 1 and 2), the effect of truncation should not be too problematic as we have a seven-year lead time before 2000.

[12] Our novelty measure reflects the *nicheness* of the technological invention publicized by the patent. It can also be interpreted as an inverse measure of *crowdedness* in the market for the technological category.

[13] Approximately 40,000 IPC subgroups are active in each period, and a single primary IPC subgroup is assigned to each patent. Refer to Section 5.1 for the details.

[14] A similar result is obtained for the transition from periods 1 to 2.

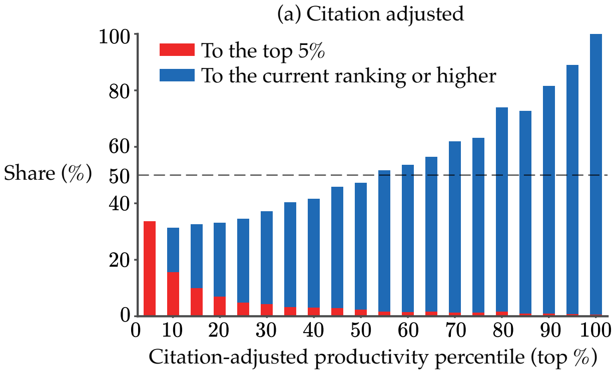


Figure 2.1: Change in the productivity class of inventors from period 0 to period 1

As discussed in the Introduction, a major reason for this downward pressure may be the obsolescence and imitations of technologies as well as decreasing returns in learning-by-doing from the extant technologies. Yet, we find that some top inventors stay highly productive, whereas some inferior ones surpass superior inventors. Each red bar in Figure 2.1 indicates the share of inventors in the corresponding class in period 0 who transitioned to the top 5% class in period 1. The transitions are observed from a wide range of lower classes.[15]

2.3 Collaborator recombinations and technological shifts

The data indicate a key relation among productivity, collaboration, and the technological specialization of inventors for knowledge creation that suggests the relevance of the BF model. Let

$$N_{it} \equiv \cup_{j \in \mathcal{G}_{it}} G_j \setminus \{i\} \tag{2.2}$$

represent the set of collaborators of inventor $i \in I$ such that each inventor in N_{it} participates in at least one common patent with i in period t. The *collaborator recombination* of inventor i in period t is then defined by the number of new collaborators in period t:[16]

$$\Delta n_{it} \equiv |N_{it} \setminus N_{i,t-1}|. \tag{2.3}$$

The average values of Δn_{it} for inventors in I are 9.84 and 6.37 in periods 1 and 2, respectively. These values coincide with the average numbers of collaborators that were replaced provided that the number of collaborators is the same across periods.

Next, define the *technological specialization* of inventor i in period t by set S_{it} of the IPC subgroups attached to the patents in which i is involved in period t. The *technological*

[15]A similar observation was made for US data between 1880 and 1940 by Akcigit et al. (2017). They found evidence that new inventors receive more patent citations than incumbent inventors.

[16]Alternatively, it may be defined by the sum of the number of new collaborations and that of separations from the collaborations in the previous period, i.e., $\Delta n_{it} = |N_{it} \setminus N_{i,t-1}| + |N_{i,t-1} \setminus N_{it}|$. The qualitative result remains the same under both definitions.

shift of inventor i is then defined, similarly to the collaborator recombination in (2.3), by the number of IPC subgroups in which i is newly specialized in period t:

$$\Delta s_{it} \equiv |S_{it} \backslash S_{i,t-1}|. \tag{2.4}$$

The average values of Δs_{it} are 4.41 and 2.66 in periods 1 and 2, respectively. High correlations, 0.55 and 0.54, between $\ln \Delta n_{it}$ and $\ln \Delta s_{it}$ in periods 1 and 2, respectively, suggest that new collaborations result in a shift of inventors' technological expertise. High correlations, 0.30 and 0.29, between the novelty-adjusted $\ln \bar{y}_{it}$ and $\ln \Delta s_{it}$ in periods 1 and 2, respectively, further indicate that technological shifts lead to higher novelty of invention.

These high correlations naturally extend to include citation-adjusted productivity. This is particularly transparent when we focus on the set of inventors in each given citation-adjusted productivity class $x = 5, 10, \ldots, 100$ persistently in both periods 1 and 2, i.e., $\Gamma(x) \equiv \cap_{t=1,2} \Gamma_t(x)$, so that inventors in $\Gamma(x)$ are persistently more productive than those in $\Gamma(x')$ for $x < x'$ for both periods.

For an inventor in class x in period t, denote *the average collaborator recombination*, *average technological shift*, and *average productivity* by $\Delta n_t(x) \equiv \frac{1}{|\Gamma(x)|} \sum_{i \in \Gamma(x)} \Delta n_{it}$, $\Delta s_t(x) \equiv \frac{1}{|\Gamma(x)|} \sum_{i \in \Gamma(x)} \Delta s_{it}$, and $\bar{y}_t \equiv \frac{1}{|\Gamma(x)|} \sum_{i \in \Gamma(x)} \bar{y}_{it}$, respectively. Figure 2.2 plots $\Delta n_t(x)$, $\Delta s_t(x)$, and novelty-adjusted $\bar{y}_t$ for $t = 1, 2$ for each citation-adjusted productivity class $x = 5, 10, \ldots, 100$, in which we find the following:

Observation 2 (Recombinations, technological shifts, and inventor productivities) *A more citation-wise productive inventor practices a more active recombination of collaborators and is associated with a larger technological shift as well as higher novelty in the created knowledge on average.*

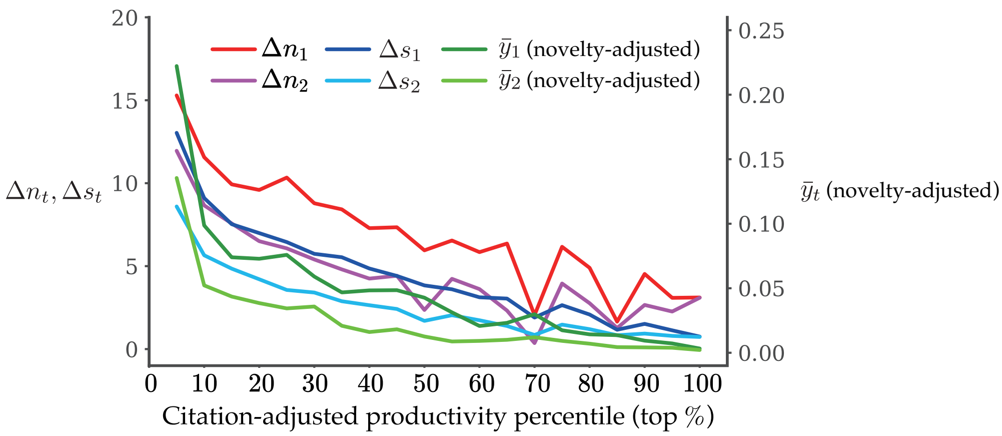


Figure 2.2: Recombinations, technological shifts, and productivities of inventors

2.4 Invention strategies and research experience

Our final observation concerns the difference in invention strategies of inventors with different research experience. To see this, let the *cumulative research scope* of inventor i in period t be quantified by the cumulative number of technological categories $k_{it} = \left|\bigcup_{t'<t} S_{it'}\right|$ such that inventor i has worked in the past. Here, a larger research scope of an inventor can be interpreted to indicate their larger research experience and, in turn, is expected to be positively correlated with their knowledge stock.[17]

Research experience naturally translates to productivity. For example, the top 5% of inventors citation-wise have, on average, 3.3 and 2.3 times larger research scope than the bottom 5% of inventors in periods 1 and 2, respectively; thus, the former can potentially rely more on their past research experience to create new knowledge than the latter.

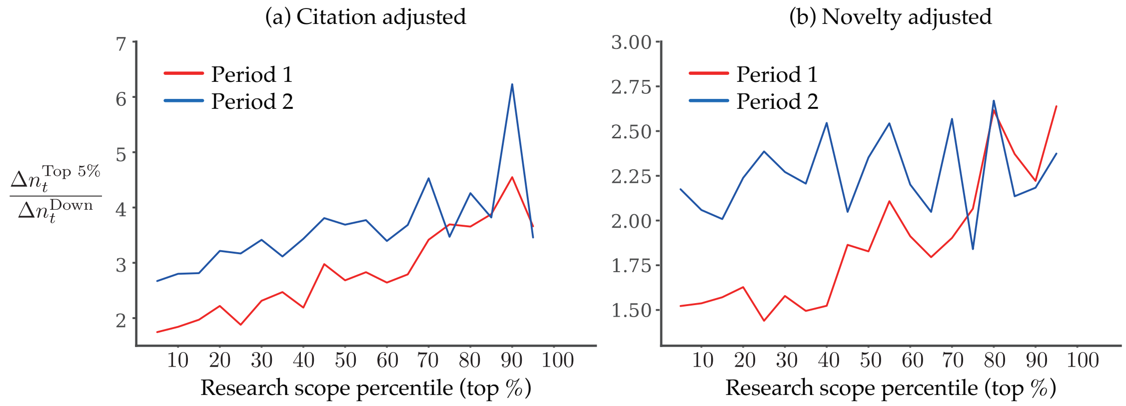


Figure 2.3: Recombination of upgrading versus downgrading inventors in period 1

Let $\Delta n_t^{\text{Top 5\%}}(x)$ represent the average size of collaborator recombinations by inventors who upgraded their productivity class from x in period $t-1$ to the top 5% in period t. Similarly, let $\Delta n_t^{\text{Down}}(x)$ be the average size of collaborator recombinations of inventors who downgraded their class from x to $x' > x$ for $x = 5, 10, \ldots, 95$.[18] Panels (a) and (b) in Figure 2.3 plot $\Delta n_t^{\text{Top 5\%}}(x)/\Delta n_t^{\text{Down}}(x)$ under citation- and novelty-adjusted productivities, respectively, against research scope percentile.[19]

Although the case of novelty-adjusted productivities in period 2 shown in Figure 2.3(b) is an exception, one can find a general tendency summarized as follows.

Observation 3 (Collaborator recombination versus stock of knowledge) *More experienced inventors with a larger research scope rely relatively more on their own stock of knowledge*

[17]Cumulative inventor productivity may be a more natural representation of knowledge stock. However, the cumulative research scope of an inventor defined in this way is less sensitive to the truncation of data due to the short time horizon because inventors stick to their past research fields due to the learning-by-doing effect (Horii, 2012). Because these two measures are expected to be positively correlated, the latter is considered to be a reasonable proxy for the former.

[18]The lowest class $x = 100$ is omitted because there is no further downgrading from there.

[19]The inventors are binned in 1 of the 20 5% interval bins of research scope percentile.

than knowledge from new collaborators for inventing, whereas the opposite is true for less experienced inventors with a smaller research scope.

3 The Berliant–Fujita model

This section briefly describes the BF model. Each agent develops new knowledge either in isolation or by collaborating in pairs, building on their accumulated stock of knowledge. Let I be the set of all agents who engage in knowledge creation, where all agents are assumed to be symmetric. Let $\delta_{ij} \in [0,1]$ be the proportion of time that agent $i \in I$ allocates for collaboration with $j \in I$. If agent i works in isolation, then their knowledge creation is subject to constant returns technology as given by $y_{ii} = \delta_{ii} a k_{ii}$ if $\delta_{ii} \in (0,1]$ and 0 otherwise, where $a > 0$, k_{ii} is the knowledge stock of agent i and y_{ii} is the output. If the subject instead collaborates with agent $j\,(\neq i)$, then their joint output y_{ij} is given by

$$y_{ij} = \delta_{ij} b \left(k_{ij}^{C}\right)^{\theta} \left(k_{ij}^{D}\right)^{\frac{1-\theta}{2}} \left(k_{ji}^{D}\right)^{\frac{1-\theta}{2}} \tag{3.1}$$

for $\delta_{ij} \in (0,1]$ and $y_{ij} = 0$ otherwise, where $b > 0$, k_{ij}^{C} is the common knowledge of i and j; k_{ij}^{D} is the knowledge of agent i differentiated from that of j; and $\theta \in (0,1)$ is the relative importance of common knowledge.

All knowledge stocks are symmetric, and the output from the collaboration of agents i and j becomes their common knowledge. Thus, the common knowledge of i and j increases as their collaboration lasts longer, and the differentiated knowledge between i with other agents also increases relative to their common knowledge. To achieve the best combination of common and differentiated knowledge with collaborators, agents collectively decide the group of collaborators, where each agent i optimally chooses δ_{ij} for each j of their collaborators to maximize the total output $\sum_j y_{ij}/2$ (assuming an equal split of output between collaborators).[20]

4 Regression model

This section introduces three regression models to identify the causal relation among the productivity of inventions, collaborators' differentiated knowledge, and the magnitude of collaborator recombination based on the BF model. In the regressions, we focus on collaborative inventions and do not address the choice between working in collaboration and working in isolation. Accordingly, our formulation assumes a strictly positive number of collaborators for each inventor in each period.

Let $t = 0, 1, \ldots, T$ be the consecutive periods in which data are available, and let I_t be the set of all inventors who participated in the development of at least one patent in

[20] Myopic core is adopted as the equilibrium concept.

period t. The subset of inventors, each of whom is involved in the development of at least one patent in every period (introduced in Section 2.1), is denoted by $I\,(\subset I_t)$.

Let $\mathcal{G}_t$ represent the set of all patents applied in period t. We call the development of each patent $j \in \mathcal{G}_t$ a *project* j. Then, G_j introduced in Section 2.1 represents the set of inventors who participated in project j. The set of projects in which inventor $i \in I_t$ participated (also introduced in Section 2.1) can be rewritten as $\mathcal{G}_{it} \equiv \{j \in \mathcal{G}_t\ :\ i \in G_j\}$. Accordingly, set N_{it} of the collaborators of inventor i in period t is given by (2.2) , and output $\bar{y}_{it}$ of inventor i is given by (2.1) in sections 2.3 and 2.1, respectively.

4.1 The Berliant–Fujita model of collaborative knowledge creation

To bring the theory to the data, we simplify the original specification and focus on the role of knowledge exchange among collaborators. First, we formulate a regression model for the knowledge creation function between an inventor and their average collaborator rather than for the total output by an inventor, thereby abstracting from the role of the number of collaborators:

$$\ln y_{it} = \alpha + \beta \ln k_{it}^D + \gamma_1 \ln k_{it} + \gamma_2 (\ln k_{it})^2 + \ln A_{it} + \lambda_i + \tau_t + \varepsilon_{it} \tag{4.1}$$

where $y_{it} \equiv \bar{y}_{it}/n_{it}$ represents the *average pairwise output* by inventor i in period t and $n_{it} \equiv |N_{it}|$. In this formulation, each piece of knowledge is assumed to be created by pairwise collaboration (even though a patent is developed by more than two collaborators), so that the time spent with an average collaborator j is assumed to be given by $\delta_{ij} = 1/n_{it}$ of period t.

Second, among the endogenous variables of the theoretical model, we focus on the differentiated knowledge k_{ji}^D of collaborators in (3.1) because this is the primary source of new ideas as discussed in Section 2 while abstracting from the role of common knowledge k_{ij}^C and differentiated knowledge k_{ij}^D of inventor i. This key variable appears as k_{it}^D in the second term on the right hand side (RHS) of (4.1) in the form of *the average differentiated knowledge of collaborators of* i and is defined by the average output that the collaborators of i produced outside the joint projects with i:

$$k_{it}^D = \frac{1}{n_{it}} \sum_{j \in N_{it}} \sum_{k \in \mathcal{G}_{jt} \setminus \mathcal{G}_{it}} \frac{g_k}{|G_k|}. \tag{4.2}$$

Here, k_{it}^D includes only the *fresh* knowledge of collaborators that they create with inventors other than i in the current period and not their knowledge stock from the past. This definition mitigates the discrepancy between theory and reality by reflecting on Observation 1 in Section 2.2 that past knowledge is strongly associated with negative effects. Note that there are no negative effects of past knowledge in the BF model because all the (infinite number of) potential pieces of knowledge are symmetric. Instead of attempting to disentangle the positive effect from the negative one in the past knowledge, we choose a simpler specification in which knowledge as a source of new ideas fully depreciates in one period.

The value of k_{it}^{D} may also be interpreted as the average productivity of i's collaborators outside the joint projects with i. This feature plays a role when we construct an instrument for this variable in Section 6.

Third, we control for the (*cumulative*) *research scope* of inventor i's past projects:

$$k_{it} = \Big|\bigcup_{t'<t} S_{it'}\Big|. \tag{4.3}$$

Because the specialized research fields of inventors persist due to the positive learning-by-doing effect (e.g., Horii, 2012), the cumulative set of technological categories of past patents in which an inventor has been involved is less sensitive to the truncation of data due to its short time horizon (as in our case) than their cumulative output in terms of patents. Nonetheless, the research scope defined this way can be considered as a proxy for the knowledge stock of i because they are naturally expected to be positively correlated.

Moreover, because the common knowledge of i and their collaborators as well as the differentiated knowledge of i are contained in the knowledge stock of i, their effects are also partly controlled for by the research scope of i. The research scope of i is also expected to control for other effects, including obsolescence, imitations, and learning-by-doing effects, on the extant technologies discussed in Section 2.2. The third and fourth terms on the RHS of (4.1) capture their overall effects up to the second order.

Finally, in the fifth term, A_{it} bundles the inventor- and time-specific productivity shifters for inventor i, $A_{it} \equiv e^{\boldsymbol{X}_{it}'\boldsymbol{\eta}}$, where $\boldsymbol{X}_{it}$ represents a vector that includes spillover effects from other inventors in the geographical neighborhood, proximity to R&D expenditure, manufacturing employment/production, and residential population. The last three terms, λ_i, τ_t, and ε_{it}, on the RHS are the time-invariant inventor fixed effect, period fixed effect, and inventor- and period-specific error, respectively. The values of parameters $\alpha, \beta, \gamma_1, \gamma_2, \boldsymbol{\eta}$, and τ_t are estimated by regressions.

4.2 Quantity–quality decomposition

The definition of inventor productivity given by (2.1) implies the log-linear relation between quantity and quality of their output:

$$\ln y_{it} = \ln y_{it}^{p} + \ln y_{it}^{q}. \tag{4.4}$$

In the first term on the RHS of (4.4), y^p_{it} denotes the quantity, i.e., the average count of patents of inventor i's pairwise output given by $y^p_{it} \equiv \bar{y}^p_{it}/n_{it}$, where $\bar{y}^p_{it} \equiv \sum_{j\in\mathcal{G}_{it}} 1/|G_j|$, which coincides with $\bar{y}_{it}$ under $g_j = 1$ in (2.1). In the second term, y^q_{it} represents the average quality of i's pairwise output $y^q_{it} \equiv y_{it}/y^p_{it}$ $\left(= \bar{y}_{it}/\bar{y}^p_{it}\right)$. We can thus decompose the effect of each explanatory variable in (4.1) into those on the quantity and quality of inventors' pairwise output y_{it} by estimating the model:

$$\ln y^m_{it} = \alpha^m + \beta^m \ln k^D_{it} + \gamma^m_1 \ln k_{it} + \gamma^m_2 (\ln k_{it})^2 + \ln A^m_{it} + \lambda^m_i + \tau^m_t + \varepsilon^m_{it} \tag{4.5}$$

for $m = p$ and q, where the coefficients of each explanatory variable for $m = p$ and q add up to those of the corresponding variable in (4.1). In particular, we have $\beta = \beta^p + \beta^q$ for the effect of collaborators' differentiated knowledge.

4.3 Recombinations and differentiated knowledge of collaborators

Our third regression model identifies the factors that determine the value of k^D_{it}:

$$\ln k^D_{it} = \tilde{\alpha} + \tilde{\beta} \ln \Delta n_{it} + \tilde{\gamma}_1 \ln k_{it} + \tilde{\gamma}_2 (\ln k_{it})^2 + \ln \tilde{A}_{it} + \tilde{\lambda}_i + \tilde{\tau}_t + \epsilon_{it} \tag{4.6}$$

where Δn_{it} given by (2.3) is considered to be endogenous, reflecting the active efforts of an inventor to find suitable collaborators.

The aim of this regression is twofold. One is to see the role of collaborator recombination in obtaining differentiated knowledge from collaborators. While the choice of the magnitude of collaborator recombination in the study by Berliant and Fujita (2008) is expressed in terms of the choice of δ_{ij} (or equivalently n_{it}) in a given period t, its consequence on the amount of differentiated knowledge of collaborators cannot be captured in our data. Thus, in (4.6), we focus on the inter-temporal recombination of collaborators to see if more substantial recombination results in a larger differentiated knowledge from collaborators on average. If so, this contributes to inventor productivity in addition to the regression results of (4.1), as indicated in Observation 2. To be consistent with the BF model, the overall effect of Δn_{it} through k^D_{it} on y_{it} is expected to exhibit positive but decreasing returns because a successful collaboration requires a certain share of common knowledge.

The other is to see if the invention strategy adopted by an inventor differs depending on their past research experience as suggested by Observation 3; however, such distinction is abstracted in the BF model.[21]

[21] We present the results of two separate estimations for (4.1) and (4.6) rather than incorporating the collaborator recombination explicitly in the estimation of (4.1). The practical reason for the separation is the estimation problem in combining the two models (refer to footnote 37 in Section 7).

5 Data

This section describes our dataset, while the details are relegated to Appendix A.

5.1 Patent data

The patent data are taken from the *published patent applications* of Japan (Artificial Life Laboratory, Inc., 2018). It provides information about published patent applications to be examined for approval rather than approved patents. The advantage of using applied patents rather than approved ones is that the flow of the former at a given point in time reflects the amount of research activity at that point more precisely than the flow of the latter. In this data, each inventor is uniquely identified as long as their name and establishment affiliation have not changed.

***Patent projects* –** Our analysis targets inventors who participated in the development of patents applied between 1995 and 2009. Because patent development typically takes several years, the productivity of an inventor is evaluated by their output over five years.[22] The choice of a five-year window also reflects the availability of census data. As described in Section 2.1, we obtain a panel with three periods: 0, 1, and 2.

Because k_{it} and Δn_{it} require information from the previous period, period 0 is not included in the regressions. The information in 2010–2016 data is used to explain the time lag between the applied and published dates as well as to count the forward citations for each patent. Consequently, our panel for regressions comprises two periods, 1 and 2 (summarized in Table A.1 in Appendix A).

We focus on the ($|I| =$) 107,724 inventors who have been active in all three periods although the information on other inventors is still used as long as they collaborated with the selected inventors. Approximately 90% of inventors in I have at least one collaborator, which justifies our focus on collaborative knowledge creation. The number of inventors in a project throughout the study period is about two on average, and the number of collaborators for an inventor is six to nine on average. This is consistent with the assumption of polyadic pairwise collaboration in the BF model.

***IPC* –** Each applied patent is associated with at least one technological classification based on the IPC, which is maintained by the World Intellectual Property Organization.[23] The IPC hierarchically classifies technologies into eight classes, subclasses, groups and then to about 40,000 subgroups.[24] The IPC's labeling scheme is consistent over time, and a newly introduced category is basically associated with a new technology. Hence, the set of technological categories specified in the IPC at a given point in

[22]The same time window is adopted by Akcigit et al. (2018).
[23]Website: http://www.wipo.int/portal/en/index.html.
[24]See Appendix A.2 for the details.

time roughly represents the set of the state-of-the-art technologies at that time, making it an appropriate proxy for the set of technological knowledge.

Although an applicant can claim more than one IPC category for their patent, we adopt only the primary IPC category of each patent to represent its technological category to avoid subjective variation. We adopt the finest subgroup classifications for our baseline analyses and the next finest subclass classifications for the robustness check, which together comprise 40,691 and 609 (38,339 and 616) categories associated with the applied patents in our data in period 1 (period 2), respectively.

Let S denote the set of all technological categories (in terms of either one of the IPC subclasses and subgroups), and the technological category assigned to patent j is $s_j \in S$. The *technological specialization* of inventor i in period t is then defined by

$$S_{it} = \cup_{j \in \mathcal{G}_{it}} \{s_j\}. \tag{5.1}$$

We control for the IPC class fixed effect to explain the possible incompatibility of the quality adjustment of patents across different technology categories, where each inventor is associated with their most frequently engaging IPC class.

5.2 Productivity and differentiated knowledge

One of our preferred measures of patent quality is the cited count following the studies by Trajitenberg (2002) and Akcigit et al. (2018).[25] In our baseline analysis, we count the forward citations of each patent within three years of publication following the study by Akcigit et al. (2018).[26] In our data, the cited counts in the first three years from publication explain more than 75% of the total cited count in the first 10 years for all samples; hence, a three-year period appears to be sufficient to evaluate the patent quality. Alternatively, patent quality is measured by its novelty within the associated IPC category as described in Section 2.1. Cultivating novel technology requires knowledge, suggesting a more direct relation between knowledge input and the novelty of inventions (see Table A.2 in Appendix A.1 for the descriptive statistics for productivity variables).[27]

[25]Cited counts may not be an optimal measure of patent quality when there is an incentive to block follow-up patents as discussed by Abrams et al. (2013).

[26]It is assumed that there is at least one self-citation, namely, $g_j \geq 1$, under the citation-adjusted measure. That is, the cited count for each patent is inflated by 1 if there is no self-citation. Some authors (e.g., Inoue et al., 2015) have argued that the citation-adjusted output of a patent project should exclude self-citations by inventors in the project. Our analyses, however, include them because there is no clear incentive to inflate the cited counts for patents (unlike the case of academic papers); hence, self-citations reflect genuine technological dependence. In fact, no qualitative difference is found between the results with and without citation weights (see Section 8.3).

[27]Appendix E.1 explores alternative specifications of patent quality for robustness check. We consider the five-year window for the citation-adjusted measure, and the IPC subclass instead of subclass is adopted for the novelty-adjusted measure. Furthermore, two more alternative measures are considered:

5.3 Locational factors

Existing studies have suggested the possible influence of various exogenous locational factors on knowledge creation. We briefly describe each factor included in the regression, with the precise definitions relegated to Appendix A.3.

R&D activities are disproportionately concentrated in large cities (see Figure A.1). If an *urban agglomeration* (UA) is defined as a contiguous area of population density of at least 1000/km^2 with the total population of at least 10,000,[28] in 2000, 99% of all inventors concentrated in UAs, 81% in the largest three UAs, and 54% in the largest UA (Tokyo). The corresponding shares for the population were 75%, 54%, and 32%, respectively. Inventors located within a 10 km buffer of any of the 453 UAs are assigned to the closest UA; otherwise, their locations are considered to be rural. In the regressions, the standard errors are clustered by UAs.[29]

Local concentrations of five types of activities are controlled: the concentrations of inventors (a_{it}^{INV}), R&D expenditure ($a_{it}^{\mathrm{R\&D}}$), manufacturing employment ($a_{it}^{\mathrm{MNF}_e}$), output ($a_{it}^{\mathrm{MNF}_o}$), and residential population (a_{it}^{POP}). Each local concentration is defined by the size of concentration in a circle of given radius around inventor i.

6 Identification by instrumental variables

This section presents our strategy for identifying the causalities of knowledge creation by dealing with the endogeneity of differentiated knowledge and recombinations of collaborators for individual inventors. There are two sources of endogeneity. One results from inventors' endogenous collaboration, i.e., network endogeneity, where unobservable influences exist on inventors' collaboration decisions and their productivities. The other results from the mutual dependence of productivities between an inventor and their collaborators through k_{it}^{D} in (4.1) (as well as (4.5)). This is the so-called "reflection problem" in the context of econometric network analysis (e.g., Manski, 1993; Bramoullé et al., 2009). While we lack any useful exogenous variations that can be used to identify the causal effects of knowledge creation (cf., e.g., Bramoullé et al., 2009; Azoulay et al., 2010; Waldinger, 2010, 2012), we argue that the endogenous variables k_{it}^{D} in (4.1) and Δn_{it} in (4.6) for inventor i can be instrumented by the average value of the same variable for the distant indirect collaborators of i.

Below, we formally define the instruments for the endogenous variables in Section

one is based on the count of patent claims, and the other is unweighted count of patents.

[28]Population data are obtained from the Population Census (2010a) by MIAC.

[29]As UAs spatially expand over time on average, we use the boundaries of UAs in 2010, each of which provides the largest spatial extent during the study period 1995–2009 on average. However, the choice of the particular time point should not affect the basic results because most inventors are concentrated in relatively large UAs whose spatial coverage is relatively stable over the study period.

6.1 and explain their exogeneity and relevance in sections 6.2 and 6.3, respectively.[30]

6.1 Instruments

Let $\bar{N}_{it}^{\ell}$ be the set of the 0-th to ℓ-th indirect collaborators of inventor i given by

$$\bar{N}_{it}^{\ell} = \bar{N}_{it}^{\ell-1} \cup \left[\cup_{j \in \bar{N}_{it}^{\ell-1}} N_{jt} \right] \quad \ell = 1, 2, \dots \tag{6.1}$$

where the set of the "0-th indirect collaborators" is defined by the set of inventors comprising i and their direct collaborators $\bar{N}_{it}^{0} \equiv N_{it} \cup \{i\}$. To obtain $\bar{N}_{it}^{\ell}$ from $\bar{N}_{it}^{\ell-1}$ for each $\ell = 1, 2, \dots$, we expand $\bar{N}_{it}^{\ell-1}$ by the union of all the direct collaborators of $j \in \bar{N}_{it}^{\ell-1}$ as in (6.1). The set of the *ℓ-th indirect collaborators* of i can then be given by

$$N_{it}^{\ell} = \bar{N}_{it}^{\ell} \backslash \bar{N}_{it}^{\ell-1} \quad l = 1, 2, \dots \tag{6.2}$$

The instruments $k_{it}^{\mathrm{IV}_{\ell}}$ for k_{it}^{D} and $\Delta n_{it}^{\mathrm{IV}_{\ell}}$ for Δn_{it} are constructed as the average values of the differentiated knowledge of collaborators and of collaborator recombination, respectively, for each ℓ-th indirect collaborator $j \in N_{it}^{\ell}$:

$$k_{it}^{D,\mathrm{IV}_{\ell}} = \frac{1}{n_{it}^{\ell}} \sum_{j \in N_{it}^{\ell}} k_{jt}^{D} \quad \text{and} \quad \Delta n_{it}^{\mathrm{IV}_{\ell}} = \frac{1}{n_{it}^{\ell}} \sum_{j \in N_{it}^{\ell}} \Delta n_{jt} \tag{6.3}$$

where $n_{it}^{\ell} \equiv \left| N_{it}^{\ell} \right|$. Alternatively, to strengthen the relevance of instruments, they may allow for the repeated appearance of the same indirect collaborators:

$$k_{it}^{D,\mathrm{IV}_{\ell}} = \frac{1}{\tilde{n}_{it}^{\ell}} \sum_{l \in N_{it}^{\ell-1}} \sum_{j \in N_{l}} k_{jt}^{D} \quad \text{and} \quad \Delta n_{it}^{\mathrm{IV}_{\ell}} = \frac{1}{\tilde{n}_{it}^{\ell}} \sum_{l \in N_{it}^{\ell-1}} \sum_{j \in N_{l}} \Delta n_{jt} \tag{6.4}$$

where $\tilde{n}_{it}^{\ell} \equiv \sum_{j \in N_{it}^{\ell-1}} n_{j}$.

6.2 Exogeneity

This section explains how our instruments can virtually eliminate the endogeneities caused by the reflection problem and inventors' unobserved variables that induce endogenous collaboration.

[30] In Appendix B, we briefly discuss the similarities and differences in the nature of endogeneity and the approach to the issue between our model and the linear-in-means models of social interactions as in the study by Bramoullé et al. (2009).

6.2.1 Reflection problem

Existing models of social interactions (e.g., Bramoullé et al., 2009; De Giorgi et al., 2010; Calvó-Armengol et al., 2009) have suggested two reasons that reflection effects in our context can be reduced by using instruments constructed from farther indirect collaborators. One is the *distance effect* such that the farther an indirect collaborator is from an inventor in the collaboration network, the smaller the influence of their output on the inventor's productivity through more distant indirect linkages on the network.[31] Thus, the reflection effects can be virtually eliminated by constructing instruments from sufficiently distant indirect collaborators. The other is the *averaging effect*. As long as the number of ℓ-th indirect collaborators increases as ℓ increases, the reflection effect on an inventor from each of their ℓ-th indirect collaborators is mitigated by averaging over a larger number of indirect collaborators provided that the effects are uncorrelated among them.[32]

Fortunately, the research network in our data comprises a set of large network components so that we could identify up to the fifth indirect collaborator for a large number of inventors. Column 1 of Table 6.1 lists the average number of the ℓ-th indirect collaborators of an inventor for $\ell = 0$ to 5, where the 0-th indirect collaborators are the direct collaborators. The number of indirect collaborators of an inventor dramatically increases from 8.52 to 4,251 (6.32 to 2,563) for $\ell = 0$ to 5 in period 1 (period 2). This suggests that the reflection emanating from each fifth indirect collaborator has only marginal effects due to distance and averaging effects.

6.2.2 Unobserved factors

We suppose that inventors with similar (observable and unobservable) characteristics have proclivities to collaborate with each other; hence, they might influence their mutual productivities. We also suppose that more distant indirect collaborators share less common characteristics with each other. Thus, we can eliminate the effects from unobserved factors by constructing instruments from sufficiently distant indirect collaborators.

The most plausible situation in which unobserved factors become problematic may arise when inventors have similar technological specialization. In this case, these inventors likely share opportunities and environment to exchange and learn ideas from each other through seminars, conferences, and journals of common research subjects, thereby affecting their R&D productivities. Our data indicate, however, that

[31] For example, in eq. (6) in the study by Bramoullé et al. (2009), the endogenous peer effect from the ℓ-th indirect peer is given by $\beta^{1+\ell}$, where $\beta \in (0,1)$ and $\ell = 0,1,2,\ldots$ with the 0-th indirect peer being the direct peer. The peer effect $\beta^{1+\ell}$ from the ℓ-th indirect peers diminishes as ℓ increases.

[32] The network component of an inventor may be influenced by the unobserved factors specific to its associated firms and establishments. This possibility will be examined in sections 8.1 and 8.2.

the commonality of research subjects between a pair of inventors diminishes rapidly and eventually becomes negligible as the degree of separation in the network increases between the pair.

Columns 2–5 in Table 6.1 list the *total research scope* $\bar{S}^{\ell}_{it} \equiv |\cup_{j \in N^{\ell}_{it}} S_{jt}|$ of i's ℓ-th indirect collaborators ($\ell = 0, 1, \ldots, 5$) in period t in terms of IPC sections, classes, subclasses, and subgroups, respectively. While an inventor, on average, specializes in 1.81, 2.47, 2.98, and 5.47 (1.53, 1.92, 2.24, and 3.71) in these categories, respectively, in period 1 (period 2) (refer to rows 12–15 in Table A.1 in Appendix A), the total research scope increases for more indirect collaborators. For their fifth indirect collaborators, these numbers reach 7.50, 83.07, 275.8, and 1400 (7.29, 70.14, 213.3, and 1007), respectively, in period 1 (period 2) (rows 6 and 12 of columns 2–5 in Table 6.1). Because the total number of IPC sections is eight, they are almost fully covered. For IPC classes, subclasses, and subgroups, they cover 98.8%, 96.9%, and 64.6% (97.9%, 94.2%, and 60.6%), respectively, among all patents applied in period 1 (period 2). Thus, the set of the fifth indirect collaborators of an inventor comprises an almost full set of specialists.

Table 6.1: Diversity and similarity of technological specialization of inventors

Indirectness ℓ	(1) Count	(2) Total research scope $\bar{S}^{\ell}_{it}$	(3)	(4)	(5)	(6) Similarity in technological specialization j^{ℓ}_{it}	(7)	(8)	(9)
		Section	Class	Subclass	Subgroup	Section	Class	Subclass	Subgroup
					Period 1				
(1) 0	8.518	3.207	7.214	10.88	27.17	0.709	0.569	0.499	0.330
	(9.321)	(1.681)	(6.313)	(10.97)	(29.65)	(0.197)	(0.229)	(0.236)	(0.214)
(2) 1	51.36	4.670	17.38	32.49	94.83	0.567	0.370	0.280	0.087
	(64.79)	(2.057)	(14.62)	(32.57)	(102.82)	(0.215)	(0.228)	(0.220)	(0.120)
(3) 2	205.41	5.860	32.48	72.72	242.03	0.493	0.280	0.193	0.045
	(284.89)	(2.074)	(23.65	(64.70)	(240.91)	(0.202)	(0.202)	(0.186)	(0.083)
(4) 3	665.27	6.691	50.07	129.69	492.17	0.433	0.213	0.135	0.025
	(944.31)	(1.869)	(30.23)	(97.96)	(433.84)	(0.188)	(0.176)	(0.155)	(0.062)
(5) 4	1794.44	7.200	67.56	199.24	866.81	0.380	0.161	0.092	0.014
	(2385.52)	(1.584)	(33.33)	(126.37)	(675.22)	(0.172)	(0.148)	(0.123)	(0.046)
(6) 5	4250.87	7.501	83.07	275.77	1399.86	0.341	0.125	0.064	0.009
	(5076.27)	(1.314)	(33.31)	(146.28)	(969.11)	(0.156)	(0.124)	(0.098)	(0.036)
					Period 2				
(7) 0	6.323	2.658	5.057	7.352	17.57	0.757	0.648	0.588	0.432
	(7.579)	(1.530)	(4.642)	(7.972)	(21.84)	(0.207)	(0.247)	(0.263)	(0.271)
(8) 1	36.79	4.073	12.35	22.07	63.30	0.582	0.404	0.312	0.100
	(48.06)	(2.006)	(10.85)	(23.28)	(74.30)	(0.246)	(0.264)	(0.258)	(0.149)
(9) 2	137.59	5.306	23.84	50.43	164.62	0.505	0.309	0.218	0.054
	(195.61)	(2.147)	(18.70)	(48.17)	(179.36)	(0.229)	(0.234)	(0.219)	(0.106)
(10) 3	424.14	6.256	38.60	93.40	343.90	0.443	0.237	0.153	0.030
	(642.62)	(2.022)	(25.89)	(77.71)	(341.13)	(0.210)	(0.204)	(0.182)	(0.077)
(11) 4	1115.16	6.888	54.59	148.78	617.58	0.390	0.182	0.106	0.018
	(1693.19)	(1.771)	(30.70)	(106.36)	(548.77)	(0.191)	(0.175)	(0.148)	(0.063)
(12) 5	2563.23	7.286	70.14	213.31	1006.73	0.347	0.141	0.074	0.011
	(3589.15)	(1.506)	(32.81)	(129.82)	(793.74)	(0.171)	(0.146)	(0.118)	(0.049)

Numbers are the average values, with standard deviations in parentheses.

The expanding research scope of more distant indirect collaborators of an inventor reflects the shrinking commonality in technological specialization between them and the inventor. It can be quantified by the average Jaccar index between the technological specialization S_{it} of inventor i and those of their indirect collaborators $j \in N_{it}^{\ell}$:

$$j_{it}^{\ell} = \frac{1}{n_{it}^{\ell}} \sum_{j \in N_{it}^{\ell}} \frac{|S_{it} \cap S_{jt}|}{|S_{it} \cup S_{jt}|} \in [0,1]. \tag{6.5}$$

A larger value of j_{it}^{ℓ} implies higher average similarity in technological specialization between inventor i and their ℓ-th indirect collaborators. In particular, it takes the value 0 if their specializations do not overlap (i.e., $S_{it} \cap S_{jt} = 0$ for all $j \in N_{it}^{\ell}$) and takes the value 1 if they are identical (i.e., $S_{it} = S_{jt}$ for all $j \in N_{it}^{\ell}$).

Columns 6–9 of Table 6.1 indicate the average values of j_{it}^{ℓ} in terms of IPC sections, classes, subclasses, and subgroups, respectively. These values between an inventor and their direct collaborators are, on average, 0.71, 0.57, 0.40, and 0.33 (0.76, 0.65, 0.59, and 0.43) in period 1 (period 2), respectively (rows 1 and 7 of columns 6–9 in Table 6.1). The numbers of technological categories that an inventor shares with their collaborators are, on average, 1.39, 1.53, 1.62, and 2.05 (1.26, 1.35, 1.41, and 1.71) at IPC section, class, subclass, and subgroup levels, respectively, in period 1 (period 2).

Between an inventor and their fifth indirect collaborators, however, the commonality of technological specialization is substantially smaller. The corresponding Jaccar indices reduce to 0.34, 0.13, 0.06, and 0.01 (0.35, 0.14, 0.07, and 0.01), respectively, in period 1 (period 2) (rows 6 and 12 of columns 6–9 in Table 6.1), The numbers of technological categories that an inventor shares with their fifth indirect collaborator are as small as 0.79, 0.40, 0.24, and 0.07 (0.80, 0.43, 0.25, and 0.07) on average, respectively, in period 1 (period 2). Thus, we conclude that as long as inventors are sufficiently far apart on the collaborator network, say fifth indirect collaborators, their research fields are virtually irrelevant.

There may still remain a concern that the formation of a network component is due to positive assortative matching among inventors. This invalidates the use of indirect collaborators to construct instruments for the endogenous variables. However, we will see in Section 7.1 that this is likely not the case. In addition, the firm- and location-specific effects underlying the similarity in productivity among indirect collaborators in the outcome of models (4.1) and (4.6) are controlled by inventors' fixed effects as well as various local factors. Hence, there is little concern about the endogeneity due to unobserved factors behind the productivity similarity among indirect collaborators.

In sections 8.1 and 8.2, we further explore the influence of time-varying factors of the firm or establishment to which an inventor belongs as well as that of the group of firms or establishments involved in a joint patent.

6.3 Relevance

We argue that the relevance of our instruments derives from assortative matching by productivity at the firm level, which is exogenous to individual inventors as their firm affiliation is predetermined in our data. In Section 6.3.1, we start by showing evidence for assortative matching among firms in their investment decisions in terms of their financial performance and worker productivity. In Section 6.3.2, we indicate that the pool of potential collaborators for an inventor is largely confined within a single firm or its affiliated partners. Hence, the firm-level assortative matching results in the exogenous positive correlation of inventor productivities among the collaborators, thereby justifying the construction of instruments based on indirect collaborators.

6.3.1 Positive assortative matching of firms by worker productivity

If firms with investment relations as well as firms and their workers exhibit assortative matching by productivity, then we expect the productivities of inventors in these matched firms to be positively correlated. Evidence for assortative matching between firms and workers can be found in the existing literature (e.g., Mendes et al., 2010; Bartolucci and Devicienti, 2013; Dauth et al., 2018).[33] While direct evidence for assortative matching among firms is not available in the literature, it is suggested by the financial and ownership data for Japanese firms (Tokyo Shoko Research, 2014).[34]

From 315,347 firms with financial information available in Japan in 2014, we identify 58,634 firm pairs with investment relationships. We then construct an (undirected) network of firms with each firm as a node and each firm pair with an investment relation as an edge. Table 6.2 indicates average values of Spearman's rank correlations for average wage in addition to four financial indices between a firm and its direct and indirect partners in the network. The "indirectness" is defined analogously to that of the inventor network so that value 0 indicates the direct investment relation and value $j \geq 1$ indicates the j-th indirect investment relation.

The firms with investment relations exhibit positive correlations in the listed financial indices as well as average wage of workers (row 1). While the correlation quickly diminishes for more distant indirect partners, the relatively high correlations persist up to the first indirect relation. Because the workers include inventors in these firms, inventor productivities are expected to be correlated among these firms.

[33] See, e.g., Eeckhout and Kircher (2018) for a theoretical model.

[34] There is indirect evidence in the literature for assortative matching by productivity among firms. Namely, Bettencourt et al. (2007); Gaubert (2018); Dauth et al. (2018) indicated evidence for spatial sorting of firms and workers by productivity. Nakajima et al. (2012) and Otazawa et al. (2018) revealed that firms with transaction linkages are geographically concentrated. See, for example, Mori and Turrini (2005); Behrens et al. (2014) for theoretical models of spatial sorting.

Table 6.2: Rank correlations of financial indices between firms with ownership

Indirectness	Avg. wage	VA/worker	Capital-asset ratio	Pretax profit-asset ratio	Third-party evaluation
(1) 0	0.1267	0.0923	0.1824	0.1465	0.2577
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(2) 1	0.0930	0.0416	0.0490	0.0280	0.0926
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
(3) 2	0.0260	0.0087	0.0045	0.0067	0.0132
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

(*i*) The numbers in the parentheses are p-values of two-sided tests. (*ii*) "Avg. wage" represents the average nominal wage per (regular) worker; "VA/worker" represents the value added per worker; "Capital-asset ratio" represents the owned capital to total asset ratio; "Pretax profit-asset ratio" represents the pretax profit to total asset ratio; and third-party evaluation is the score ranging in [0,100] based on over 200 financial indices provided by the Tokyo Shoko Research.

6.3.2 Collaboration and firm affiliation of inventors

If the size of a firm/establishment in period t is defined by the number of inventors who belong to the firm/establishment at some point in the period, then the average and median of the firm size are 26,923 and 7,757 (23,025 and 8,207) while those of the establishment size are 3,500 and 1,059 (2,972 and 962) in period 1 (period 2).

In Table 6.3, columns 1 and 2 indicate the average shares of the ℓ-th indirect collaborators of an inventor who belong to the same firm as the inventor ($\ell = 0,1,\dots,5$) and columns 3 and 4 show similar shares for establishments in periods 1 and 2, respectively. Note that on average, more than 80% of collaborators are confined within a single firm as well as within a single establishment.[35] Although the shares decrease as ℓ increases, they still remain as high as 25.6% and 31.7% for the fifth indirect collaborator for firms and 20.6% and 25.3% for establishments in periods 1 and 2, respectively.

To assess the number of firms involved in order to reach the ℓ-th indirect collaborators, we construct the collaboration network of firms with each firm as a node and each pair of firms conducting a joint patent development as an edge. Columns 5 and 6 of Table 6.3 list the values for the average shortest-path length from the firm of an inventor to the firm of their ℓ-th indirect collaborator on this network in periods 1 and 2, respectively. Although the shortest-path length, i.e., the smallest number of distinct firms to reach the ℓ-th indirect collaborator, increases as ℓ increases in both periods, it still remains smaller than two even for the fifth indirect collaborator.

The pool of potential collaborators for an inventor is mostly confined to a single firm or its closely affiliated firms, and is essentially exogenous to inventors in our regression as we focus on the inventors who do not change their firm affiliation. Because these affiliated firms exhibit assortative matching in terms of worker productivity, provided

[35] For the samples limited to the Japanese patents applied by corporations, more than 90% of inventions occur within a single establishment (Inoue et al., 2017).

that joint R&D occurs more often among firms with closer investment relations, the productivities of inventors in the potential pool of collaborators are expected to be positively correlated. Given that k_{it}^D is the average productivity of collaborators of i outside the joint projects with i, the values of $\ln k_{it}^D$ in the affiliated firms are expected to be positively correlated.

Table 6.3: Firm and establishment affiliations of inventors

	(1)	(2)	(3)	(4)	(5)	(6)
	Same firm share		Same establishment share		Path length to firm	
Indirectness	Period 1	Period 2	Period 1	Period 2	Period 1	Period 2
(1) 0	0.819	0.824	0.811	0.814	0.453	0.399
	(0.248)	(0.261)	(0.263)	(0.275)	(0.498)	(0.490)
(2) 1	0.721	0.731	0.687	0.694	0.887	0.789
	(0.271)	(0.289)	(0.303)	(0.319)	(0.568)	(0.626)
(3) 2	0.616	0.640	0.565	0.584	1.179	1.107
	(0.294)	(0.308)	(0.325)	(0.338)	(0.543)	(0.630)
(4) 3	0.501	0.539	0.440	0.473	1.410	1.354
	(0.302)	(0.316)	(0.328)	(0.340)	(0.534)	(0.625)
(5) 4	0.377	0.430	0.318	0.360	1.633	1.584
	(0.293)	(0.311)	(0.307)	(0.325)	(0.525)	(0.618)
(6) 5	0.256	0.317	0.206	0.253	1.843	1.794
	(0.260)	(0.290)	(0.260)	(0.291)	(0.499)	(0.602)

(*i*) Numbers in parentheses are standard deviations. (*ii*) "Same firm share" and "same establishment share" denote the shares of ℓ-th indirect collaborators ($\ell = 0, 1, \ldots, 5$) of an inventor who belong to the same firm and the same establishment as the inventor, respectively. (*iii*) "Path length to firm" means the average number of firms on the shortest path from an inventor to the ℓ-th indirect collaborator on the research collaboration network of firms.

As for Δn_{it}, recall Observation 2 in Section 2.3 that inventors with higher productivities conduct more active recombination of collaborators. Consequently, the size of the collaborator recombination Δn_{it} is expected to be similar among indirect collaborators with similar productivities. Yet, between inventor i and their indirect collaborator j, the relevance between Δn_{it} and Δn_{jt} induced by the assortative matching among firms and workers is weaker than that between k_{it}^D and k_{jt}^D because the former pair is not directly related to the productivities of i and j unlike the latter pair. Thus, rather than (6.3), we adopt the alternative instrument given by (6.4) for Δn_{it} that focuses more on indirect collaborators who are more frequently connected to i.

7 Regression results

This section presents our main regression results for models (4.1), (4.5), and (4.6). In all the regressions conducted, the fixed effects of inventors, periods, and IPC classes (see Section 5.1) are controlled. The local factors described in Section 5.3, except for residential population, are constructed for a circle with a 1 km radius around each inventor, whereas it is set to 20 km for residential population to account for urban environments. Standard errors are clustered by UAs (refer to Section 5.3).[36]

7.1 The Berliant-Fujita model

Table 7.1 summarizes the regression results for model (4.1), with columns 1–5 and 6–10 indicating the results for citation- and novelty-adjusted productivities, respectively. Columns 1 and 6 report the results from the ordinary least squares (OLS) regression for the respective cases, and the rest report those from the two-stage least squares (2SLS) instrumental variable (IV) regressions. For the IV regressions, we use the third to fifth indirect collaborators to construct IVs for $\ln k_{it}^{D}$. More specifically, we use all three instruments $\ln k_{it}^{\mathrm{IV}_\ell}$ for $\ell = 3, 4$ and 5 in column 2 (column 7), while only one of them is used in columns 3–5 (8–10), respectively, for citation (novelty)-adjusted productivity.[37] To make the results comparable, the observations are restricted to the set of 58,464 inventors (rather than the 107,724 considered in sections 2 and 5), with at least one fifth indirect collaborator.[38]

The IV results support the mechanism of knowledge creation proposed by Berliant and Fujita (2008) (row 1, columns 2–5 and 8–10). In particular, the estimated coefficients of $\ln k_{it}^{D}$ are persistently positive, 0.27–0.29 (0.34–0.38), and significant for citation- (novelty)-adjusted productivity; however, values below 1 indicate decreasing returns to the differentiated knowledge of collaborators as the benefit from collaborators' differentiated knowledge will eventually be dominated by that of common knowledge with collaborators and differentiated knowledge of the inventor.

The estimated positive effect of research scope $\ln k_{it}$ of an inventor (row 2) and the negative effect of its squared term $(\ln k_{it})^2$ (row 3) are consistent with the positive but decreasing returns of learning-by-doing from the extant technologies discussed in sections 2 and 4. However, because $\ln k_{it} > 0$ from the definition of k_{it} (≥ 1) in our data, the second-order effects appear to dominate the first-order effects; in other words, the net effect of $\ln k_{it}$ is mostly negative. The overall negative effects associated with past knowledge imply that the positive learning-by-doing effects are dominated by the negative effects from imitations and obsolescence, which accounts for the persistent downward pressure on inventor productivity in Observation 1 in Section 2.2.

[36]Because the instruments $\ln k_{it}^{D,\mathrm{IV}_\ell}$ for $\ln k_{it}^{D}$ in (4.1) and (4.5) as well as $\ln \Delta n_{it}^{\mathrm{IV}_\ell}$ for $\ln \Delta n_{it}$ in (4.6) involve inventors located in different UAs, one might suspect that cluster-robust standard errors are incorrect because the instruments for any inventor i might be correlated with errors ε_{jt} in (4.1), ε_{jt}^{m} in (4.4), and ϵ_{jt} in (4.6) for any inventor j even if inventors i and j are located in different UAs. However, we consider that these cluster-robust standard errors still provide correct standard errors because the inventor fixed effects are controlled in all regressions that encompass UA-specific fixed effects, making the errors free from the correlation with UAs while allowing for standard errors to vary across UAs.

[37]It looks as if the instrument $\ln \Delta n_{it}^{\mathrm{IV}_\ell}$ for $\ln \Delta n_{it}$ works as an instrument for $\ln k_{it}^{D}$ in the estimation of (4.1) because $\ln \Delta n_{it}^{\mathrm{IV}_\ell}$ has relevance with $\ln k_{it}^{D}$ via (4.6). However, the relevance turned out to be weak between $\Delta n_{it}^{\mathrm{IV}_\ell}$ and k_{it}^{D} although we find that Δn_{it} has positive significant effect on k_{it}^{D} in Section 7.3.

[38]The basic properties of each variable remain the same as described in Table A.1.

For all the choices of IVs, the first-stage F values are large (row 12, columns 2–5 and 7–10), meaning that the IVs do not seem to be weak.[39] To confirm the exogeneity of the IVs, we use $\ln k_{it}^{\mathrm{IV}_\ell}$ for all $\ell = 3,4$ and 5 in columns 2 and 7 for citation- and novelty-adjusted productivities, respectively, and conduct Hansen's (1982) J test for overidentifying restrictions. The p-values of the test are 0.928 and 0.768 for the respective cases (row 11, columns 2 and 7), meaning that the exogeneity of the IVs cannot be rejected.[40] Moreover, the estimated coefficients for the alternative choices of the IVs are remarkably similar (compare columns 2–5 and columns 7–10), which indicates that these IVs are reasonably exogenous.

The OLS result is consistent with the IV results in terms of the signs of the estimated coefficients, but it appears to have biases in several estimated coefficients. For the effect of $\ln k_{it}^{D}$, we find downward bias in the OLS estimates (compare columns 1 and 2–5 and columns 6 and 7–10 in row 1).[41] An explanation for the bias is that the more productive inventors attract (or are assigned by their firm) a larger number of relatively unexperienced collaborators and inherit more collaborators with lower productivity than the inventor intends. The removal of this reverse causality has lead to a larger positive selection effect in the IV estimates.

If the collaborations were subject to endogenous positive assortative matching among inventors, then the OLS estimate of β is expected to be biased upward and not downward because the positive assortative matching among inventors implies positive reflection effects (see, e.g., Bramoullé et al., 2009, for the case of social interaction models). The possible bias for this reason is, therefore, at least not dominating.

Under the OLS, this selection effect may be partly picked up by the effect of the local concentration of inventors $\ln a_{it}^{\mathrm{INV}}$, which has upward bias (compare columns 1 and 2 and columns 6 and 7 in row 4). This is because a larger inventor concentration induces more fruitful collaborations, resulting in larger average differentiated knowledge from collaborators. Consequently, in the IV result, the part of the OLS estimate of the coefficient of $\ln a_{it}^{\mathrm{INV}}$ for which the collaborator recombination is responsible is absorbed into the coefficient of $\ln k_{it}^{D}$. What is left in the estimated effect of $\ln a_{it}^{\mathrm{INV}}$ may be interpreted as the positive spillover effect from the local inventor concentration.[42] Specifically, a 10% increase in the inventor concentration results in 1.2% and 1.8–2.0% increases in citation- and novelty-adjusted productivity, respectively.

The concentration of R&D expenditure has a persistent positive effect for all the specifications (row 5), where its 10% increase raises citation- and novelty-adjusted

[39] See Table C.1 in Appendix C for the results of the first-stage regressions.

[40] Of course, this result of Hansen's J test is by no means sufficient to guarantee the exogeneity of the instruments if all the instruments are subject to the same type and magnitude of bias.

[41] Akcigit et al. (2018) reported a similar downward bias on the effects of interaction levels on innovation productivity within a patent team.

[42] In Section 8.2, we find that the effect of $\ln a_{it}^{\mathrm{INV}}$ is largely explained by factors specific to the establishments affiliated through joint patent projects.

productivities by 0.26% and 0.35–0.36%, respectively. The size of manufacturing output has essentially no impact on inventor productivity.

The positive significant effects of local manufacturing employment on citation-adjusted productivity (row 6, columns 2–5), where its 10% increase raises productivity by 0.23–0.24%, reflect that innovations are linked to production. Furthermore, citations are often made by the related production units of nearby firms. It is insignificant for novelty-adjusted productivity (row 5, columns 7–10) as technological novelty is not necessarily relevant for the present production levels.

The nearby residential population appears to be irrelevant for inventor productivity.

Table 7.1: Regression results for (4.1) (Dependent variable: $\ln y_{it}$)

	Citations					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln k^{D}_{it}$	0.163***	0.286***	0.286***	0.287***	0.273***	0.164***	0.344***	0.341***	0.353***	0.377***
	(0.0102)	(0.0254)	(0.0257)	(0.0353)	(0.0399)	(0.00493)	(0.0310)	(0.0335)	(0.0296)	(0.0629)
(2) $\ln k_{it}$	0.110***	0.0931***	0.0931***	0.0929***	0.0949***	0.147***	0.114***	0.115***	0.113***	0.108***
	(0.0153)	(0.0119)	(0.0119)	(0.0143)	(0.0165)	(0.0172)	(0.0228)	(0.0229)	(0.0226)	(0.0248)
(3) $(\ln k_{it})^2$	-0.0890***	-0.0820***	-0.0820***	-0.0820***	-0.0828***	-0.195***	-0.178***	-0.178***	-0.177***	-0.175***
	(0.00967)	(0.00868)	(0.00865)	(0.00954)	(0.00991)	(0.00926)	(0.00594)	(0.00564)	(0.00665)	(0.0108)
(4) $\ln a^{\mathrm{INV}}_{it}$	0.171***	0.117*	0.117*	0.117*	0.123**	0.310***	0.200**	0.202**	0.195**	0.180***
	(0.0579)	(0.0633)	(0.0635)	(0.0597)	(0.0540)	(0.0913)	(0.0939)	(0.0965)	(0.0887)	(0.0672)
(5) $\ln a^{\mathrm{R\&D}}_{it}$	0.0272***	0.0256***	0.0256***	0.0256***	0.0258***	0.0420***	0.0364***	0.0365***	0.0362***	0.0354***
	(0.00786)	(0.00679)	(0.00679)	(0.00664)	(0.00670)	(0.0156)	(0.0127)	(0.0128)	(0.0125)	(0.0120)
(6) $\ln a^{\mathrm{MNF}e}_{it}$	0.0149***	0.0240***	0.0240***	0.0240***	0.0230***	-0.00859	0.0132	0.0128	0.0143	0.0172
	(0.00566)	(0.00438)	(0.00436)	(0.00533)	(0.00598)	(0.0105)	(0.00989)	(0.00955)	(0.0108)	(0.0158)
(7) $\ln a^{\mathrm{MNF}o}_{it}$	0.00832	0.00522	0.00522	0.00520	0.00555	-0.00362	-0.00512	-0.00509	-0.00519	-0.00539
	(0.00581)	(0.00804)	(0.00806)	(0.00779)	(0.00732)	(0.00552)	(0.00721)	(0.00717)	(0.00732)	(0.00761)
(8) $\ln a^{\mathrm{POP}}_{it}$	-0.449	-0.660	-0.660	-0.661	-0.637	0.793*	0.0701	0.0837	0.0346	-0.0611
	(0.519)	(0.490)	(0.490)	(0.493)	(0.470)	(0.442)	(0.415)	(0.427)	(0.390)	(0.358)
(9) τ_1	0.227***	0.173***	0.173***	0.172***	0.178***	0.304***	0.173***	0.175***	0.166***	0.149***
	(0.0159)	(0.0150)	(0.0149)	(0.0213)	(0.0245)	(0.0307)	(0.0382)	(0.0403)	(0.0352)	(0.0477)
(10) R^2	0.151					0.184				
(11) Hansen J p-val.		0.928					0.768			
(12) 1st stage F		727.1	2178	1080	509.6		557.6	1590	918.7	471.4
(13) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

7.2 Quantity-quality decomposition

In this section, the effect of each explanatory variable in (4.1) is decomposed into the fractions that accrue to the quantity and quality of an inventor's output. The regression result for the former is relegated to Appendix D (Table D.1), and those for the latter are given in Table 7.2, which is organized similarly to Table 7.1.[43] Together with the results shown in Table 7.1, the present results reveal the extent to which each explanatory variable contributes to the quality and quantity of the patent output.

[43]The first stage of the regression is shared with (4.1). To confirm the exogeneity of the IVs, we use $\ln k^{\mathrm{IV}\ell}_{it}$ for all $\ell = 3,4$ and 5 in columns 2 and 7 for quality- and novelty-adjusted productivities, respectively. We conduct Hansen's (1982) J test for overidentifying restrictions. The p-values of the test are 0.419 and 0.314 for quality- and novelty-adjusted productivities, respectively (row 11, columns 2 and 7), meaning that the exogeneity of the IVs cannot be rejected.

We find contrasting roles of differentiated knowledge of collaborators between the two measures of patent quality: more than 90% of its contribution is attributed to increasing the quantity rather than quality of research output under the citation-adjusted measure (row 1 and columns 2–5 in tables 7.1 and 7.2), whereas around 65% of the contribution accrues to increasing the quality rather than quantity of research output under the novelty-adjusted measure (row 1 and columns 7–10 in tables 7.1 and 7.2).

This result indicates that knowledge exchange is an especially effective source of technological novelty. This appears to be the key factor for inducing the technological shift of an inventor to a new niche, which is consistent with the results of Berliant and Fujita (2008) as well as Horii (2012).

The decompositions of the effects of other explanatory variables are also worth explanations although there are no formal theories that account for them. For both citation- and novelty-adjusted productivity measures, the inventor as well as R&D expenditure concentrations exhibit positive significant effect on the quantity but not on the quality of inventions (rows 4 and 5 in tables 7.2 and D.1). The manufacturing employment concentration raises the quality rather than the quantity of inventions (rows 6 and 7 in tables 7.2 and D.1). The former result suggests that positive externalities from researcher agglomeration promote starting inventions, whereas the latter result may reflect that the proximity to manufacturing concentration promotes targeted inventions with higher quality.

Table 7.2: Regression results for (4.5) (Dependent variable: $\ln y_{it}^{q}$)

	Citations					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln k_{it}^{D}$	0.0273***	0.0264**	0.0269**	0.0154	0.00321	0.119***	0.230***	0.231***	0.221***	0.247***
	(0.00169)	(0.0120)	(0.0119)	(0.0192)	(0.0221)	(0.00278)	(0.0143)	(0.0146)	(0.0168)	(0.0330)
(2) $\ln k_{it}$	0.0104*	0.0106	0.0105	0.0121	0.0138*	0.0364	0.0162	0.0161	0.0179	0.0132
	(0.00578)	(0.00714)	(0.00710)	(0.00832)	(0.00758)	(0.0240)	(0.0274)	(0.0274)	(0.0272)	(0.0272)
(3) $(\ln k_{it})^2$	-0.00549***	-0.00554***	-0.00551***	-0.00616***	-0.00684***	-0.108***	-0.0976***	-0.0976***	-0.0985***	-0.0961***
	(0.00101)	(0.00138)	(0.00136)	(0.00189)	(0.00202)	(0.00520)	(0.00651)	(0.00655)	(0.00653)	(0.00657)
(4) $\ln a_{it}^{\mathrm{INV}}$	-0.0364***	-0.0361**	-0.0363**	-0.0313**	-0.0260**	0.0719**	0.00411	0.00368	0.00967	-0.00596
	(0.0126)	(0.0141)	(0.0142)	(0.0137)	(0.0102)	(0.0318)	(0.0392)	(0.0397)	(0.0383)	(0.0400)
(5) $\ln a_{it}^{\mathrm{R\&D}}$	-0.00252	-0.00251	-0.00252	-0.00236	-0.00220	0.0118*	0.00839	0.00837	0.00868	0.00789
	(0.00414)	(0.00421)	(0.00422)	(0.00409)	(0.00377)	(0.00638)	(0.00570)	(0.00569)	(0.00569)	(0.00598)
(6) $\ln a_{it}^{\mathrm{MNF}e}$	0.0271***	0.0271***	0.0271***	0.0262***	0.0254***	0.00798	0.0215**	0.0216**	0.0204**	0.0235**
	(0.00559)	(0.00521)	(0.00520)	(0.00531)	(0.00641)	(0.00850)	(0.00929)	(0.00919)	(0.00949)	(0.0112)
(7) $\ln a_{it}^{\mathrm{MNF}o}$	0.00861	0.00863	0.00862	0.00891*	0.00921*	-0.00636	-0.00728	-0.00729	-0.00720	-0.00742
	(0.00556)	(0.00534)	(0.00535)	(0.00518)	(0.00526)	(0.00458)	(0.00656)	(0.00657)	(0.00637)	(0.00684)
(8) $\ln a_{it}^{\mathrm{POP}}$	-0.582**	-0.580**	-0.581**	-0.562**	-0.541**	0.610	0.163	0.160	0.200	0.0971
	(0.238)	(0.240)	(0.239)	(0.250)	(0.251)	(0.440)	(0.477)	(0.476)	(0.479)	(0.497)
(9) τ_1	0.101***	0.102***	0.101***	0.107***	0.112***	0.151***	0.0699***	0.0694***	0.0765***	0.0578*
	(0.0180)	(0.0223)	(0.0223)	(0.0245)	(0.0200)	(0.0202)	(0.0210)	(0.0209)	(0.0225)	(0.0323)
(10) R^2	0.086					0.140				
(11) Hansen J p-val.		0.419					0.314			
(12) 1st stage F		727.1	2178	1080	509.6		557.6	1590	918.7	471.4
(13) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

The results of our regressions identify the causal relation suggested by the BF model concerning the correlation between collaborators' differentiated knowledge and inventor productivity in Observation 2 in Section 2.3.[44] In particular, technological shift Δs_{it}, which was found to be correlated with higher productivity in Observation 2, is intentionally directed toward less explored niches because of inventors' (or firms') quest for more novel invented technologies. The technological shift caused by knowledge exchange appears to be a means to overcome the negative effects of inventors' past knowledge stock as pointed out in Observation 1. This part of the result is missing in the BF model, which assumes symmetry among all pieces of knowledge. However, this finding agrees with the theoretical result of Horii (2012), who considered a more realistic economy with demand for new technologies.

7.3 Recombination and the differentiated knowledge of collaborators

This section presents the results for model (4.6), which incorporates the fundamental causality assumed in the BF model that collaborator recombination is an effective means to collect novel ideas for knowledge creation. The regression results are summarized in Table 7.3, which is organized similarly to Table 7.1 except that the dependent variable is $\ln \Delta k_{it}^{D}$ and $\ln \Delta n_{it}^{\mathrm{IV}_\ell}$ for $\ell = 3, 4$, and 5 serves as the IV for an endogenous variable $\ln \Delta n_{it}$.

In the IV results, the estimated coefficients of $\ln \Delta n_{it}$ are persistently positive 1.24–1.52 (1.71–1.96) and significant for citation- (novelty)-adjusted productivity. This confirms our earlier finding in Observation 2 on the implication from Berliant and Fujita (2008) that collaborator recombination is an effective means to acquire new ideas to facilitate invention. These estimated elasticities are greater than 1. However, because research productivity exhibits decreasing returns in the input of collaborators' differentiated knowledge, the overall effect of collaborator recombination on inventor productivity will be positive but diminishing. Specifically, putting the results from (4.1) and (4.6) together, we find that a 10% increase in collaborator recombination, presumably by utilizing the differentiated knowledge of collaborators, induces 3–4% and 6–8% increases in the respective pairwise output of an inventor.

The effect of cumulative research scope $\ln k_{it}$ in collaborator recombination in model (4.6) appears in contrast to its net negative effect on knowledge creation in model (4.1). In the IV results for (4.6), we find a positive increasing returns effect of research scope in acquiring differentiated knowledge from collaborators (columns 2–5 and 7–10 of rows 2 and 3). On the one hand, a highly established inventor with a large research scope

[44] Although we use inventor productivity $\bar{y}$ in Section 2 rather than pairwise productivity y_{it}, these are highly correlated, with correlation coefficients of 0.73 and 0.76 in periods 1 and 2, respectively. Thus, the observations made for $\bar{y}$ in Section 2 basically apply to y_{it} as well.

can attract able collaborators without a large effort, i.e., without much replacement of collaborators. On the other hand, an inventor with a small research scope must induce significant effort to find appropriate collaborators for successful inventions (or their firm should arrange so), thereby leading to a large number of new collaborators. This accounts for the mechanism behind Observation 3. While past research experience is useful to achieve better collaboration partners, it may hinder their invention as the old knowledge quickly reduces to obsolescence and is subject to imitation.

Table 7.3: Regression results for (4.6) (Dependent variable: $\ln k_{it}^{D}$)

	Citations					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln \Delta n_{it}$	0.104***	1.372***	1.400***	1.523***	1.236***	0.244***	1.722***	1.718***	1.962***	1.714***
	(0.00626)	(0.0629)	(0.0748)	(0.119)	(0.132)	(0.00792)	(0.0847)	(0.0914)	(0.157)	(0.131)
(2) $\ln k_{it}$	0.131***	-0.0220	-0.0253	-0.0401	-0.00554	0.147***	-0.0313	-0.0308	-0.0601	-0.0303
	(0.0427)	(0.0669)	(0.0653)	(0.0814)	(0.0732)	(0.0338)	(0.0638)	(0.0632)	(0.0786)	(0.0669)
(3) $(\ln k_{it})^2$	-0.0364**	0.223***	0.229***	0.254***	0.195***	-0.0422***	0.261***	0.260***	0.310***	0.259***
	(0.0156)	(0.0197)	(0.0167)	(0.0422)	(0.0392)	(0.0161)	(0.0233)	(0.0223)	(0.0476)	(0.0360)
(4) $\ln a_{it}^{\mathrm{INV}}$	0.387***	0.0138	0.00580	-0.0304	0.0539	0.515***	0.0800	0.0813	0.00957	0.0825
	(0.0916)	(0.0426)	(0.0467)	(0.0461)	(0.0507)	(0.118)	(0.103)	(0.107)	(0.0878)	(0.0939)
(5) $\ln a_{it}^{\mathrm{R\&D}}$	0.0134	0.000705	0.000432	-0.000799	0.00207	0.0320*	0.0172*	0.0172*	0.0148	0.0173*
	(0.0111)	(0.00478)	(0.00487)	(0.00571)	(0.00447)	(0.0165)	(0.00896)	(0.00895)	(0.00939)	(0.00915)
(6) $\ln a_{it}^{\mathrm{MNF}e}$	-0.0706***	-0.0139	-0.0127	-0.00720	-0.0200	-0.110***	-0.0436**	-0.0438**	-0.0329	-0.0440*
	(0.0220)	(0.0147)	(0.0151)	(0.0183)	(0.0137)	(0.0186)	(0.0219)	(0.0213)	(0.0302)	(0.0239)
(7) $\ln a_{it}^{\mathrm{MNF}o}$	0.0214	0.00814	0.00786	0.00657	0.00957	0.00221	-0.0133	-0.0132	-0.0158	-0.0132
	(0.0215)	(0.00992)	(0.0101)	(0.0110)	(0.00960)	(0.0265)	(0.00922)	(0.00923)	(0.0107)	(0.00920)
(8) $\ln a_{it}^{\mathrm{POP}}$	1.371	-0.552	-0.594	-0.780	-0.345	3.574***	1.332	1.338	0.968	1.345
	(1.043)	(1.229)	(1.217)	(1.403)	(1.273)	(1.137)	(1.050)	(1.037)	(1.247)	(1.103)
(9) τ_1	0.415***	0.514***	0.517***	0.526***	0.504***	0.698***	0.814***	0.814***	0.833***	0.814***
	(0.0269)	(0.0504)	(0.0525)	(0.0481)	(0.0428)	(0.0373)	(0.0285)	(0.0287)	(0.0310)	(0.0291)
(10) R^2	0.160					0.178				
(11) Hansen J p-val.		0.255					0.363			
(12) 1st stage F		237.7	639.9	338.5	253.9		237.7	639.9	338.5	253.9
(13) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

As other local factors appear to play minor roles, we find that the research experience and collaborator recombination remain two effective means for an inventor to improve their productivity. Taken together, a rather intricate mechanism underlying the churning of inventor productivities in Observation 1 has been disentangled and explained from the micro-level behavior of individual inventors à la Berliant and Fujita (2008) and Horii (2012).

For all the choices of IVs, the first-stage F values are large (row 12, columns 2–5 and 7–10), suggesting that the IVs are not weak.[45] To confirm the exogeneity of IVs, we use $\ln \Delta n_{it}^{\mathrm{IV}\ell}$ for all $\ell = 3, 4$ and 5 in columns 2 and 7 for citation- and novelty-adjusted productivities, respectively, and conduct Hansen's (1982) J test for overidentifying restrictions. The p-values of the test are 0.255 and 0.363 for citation- and novelty-adjusted productivities, respectively (row 11, columns 2 and 7), meaning that the exogeneity of the IVs cannot be rejected.[46] The estimated coefficients for the alternative

[45] See Table C.2 in Appendix C for the results from the first-stage regressions.

[46] The same caveat stated in footnote 40 applies here.

choices of IVs are less stable than those for model (4.1); however, they agree with each other qualitatively (columns 2–5 and columns 7–10).

The OLS estimate of the effect of $\ln \Delta n_{it}$ is consistent with the IV results in terms of the signs and significance of the estimated coefficients; however, it exhibits substantial downward bias (compare columns 1 and 2 and columns 6 and 7 in row 1). Under the OLS, a part of the effect of collaborator recombination appears in that of local inventor concentration because a larger inventor concentration implies a larger pool of potential collaborators. The downsized effect of $\ln a_{it}^{\mathrm{INV}}$ in the IV regression is consistent with this interpretation (compare columns 1 and 2–5 and columns 6 and 7–10 in row 4).

Another source of the bias is reverse causality. A higher productivity for an inventor is, on average, associated with the larger differentiated knowledge of their collaborators as well as a larger research scope. This bias appears to be reflected in the estimated coefficient of the research scope $\ln k_{it}$, which has substantial upward bias in the OLS (compare columns 1 and 2–5 and columns 6 and 7–10 in row 2).

8 Robustness

This section checks the robustness of our baseline results with emphasis on two aspects: the influence of time-varying firm and establishment-specific factors in Section 8.1 and that of time-varying affiliated firm and establishment-specific factors in Section 8.2. We investigate the sensitivity of our baseline results under alternative definitions for inventor productivities and alternative neighborhood sizes to define local factors, in addition to alternative IVs in Section 8.3.

8.1 Size and research scope of a firm and an establishment

This section considers two time-varying properties of the firm and establishment to which each inventor belongs. Let F_{it} be the set of inventors who belong to the same firm as inventor i at some point in period t, and let $F_{-i,t} \equiv F_{it} \setminus (N_{it} \cup \{i\})$, i.e., F_{it} *excluding* i *and their collaborators*.

The first property is the *firm size*, $f_{it} = |F_{-i,t}|$, representing the magnitude of the R&D activities within the firm to which inventor i belongs; however, outside the projects, an inventor and collaborators are directly involved. Given that more than 80% of collaboration occurs within a firm on average, the variation in k_{it}^{D} as well as that of Δn_{it} may simply reflect firm size in period t.[47]

[47]Note that firm size here is rather special as it aggregates all the inventors affiliated with a given firm at some point in the given period. Firm size may be slightly overstated because inventors who simply changed establishments within a firm in the same period were counted multiple times. Nonetheless, it should reflect the basic variation in the number of inventors involved in a given firm.

The second property is the *research scope* of the firm of a given inventor i defined by $s_{it}^f = |\cup_{j\in F_{it}} S_{jt} \backslash (\cup_{u\in N_{it}\cup\{i\}} S_{ut})|$, which counts the number of distinct technological categories in which patents are developed in the firm of inventor i and excludes those associated with the patents developed by i and by their collaborators. The values of f_{it} and s_{it}^f reflect the potential scale effect of a firm; for example, the availability of common research facilities, funding, and other sources of increasing returns as well as interdisciplinary spillover.

In a similar manner, we can define the set E_{it} of inventors who belong to the same establishment as inventor i in period t and set the *establishment size* $e_{it} = |E_{-it}|$ as well as the research scope $s_{it}^e = |\cup_{j\in E_{it}} S_{jt} \backslash (\cup_{u\in N_{it}\cup\{i\}} S_{ut})|$ of their establishment.

Table 8.1: Regression results for (4.1) with firm- and establishment-specific factors

	Citations				Novelty			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) $\ln k_{it}^D$	0.297***	0.271***	0.251***	0.201***	0.335***	0.276***	0.308***	0.229***
	(0.0338)	(0.0341)	(0.0420)	(0.0338)	(0.0360)	(0.0325)	(0.0429)	(0.0300)
(2) $\ln f_{it}$	-0.0370	-0.217***			0.00174	-0.276***		
	(0.0721)	(0.0572)			(0.104)	(0.107)		
(3) $\ln s_{it}^f$		0.158***				0.254***		
		(0.0220)				(0.0272)		
(4) $\ln e_{it}$			0.180***	0.0158			0.202***	-0.0351
			(0.0416)	(0.0610)			(0.0380)	(0.0893)
(5) $\ln s_{it}^e$				0.122***				0.188***
				(0.0221)				(0.0414)
(6) $\ln k_{it}$	0.0867***	0.0857***	0.0853***	0.0850***	0.103***	0.106***	0.0992***	0.102***
	(0.0210)	(0.0205)	(0.0251)	(0.0256)	(0.0239)	(0.0234)	(0.0238)	(0.0234)
(7) $(\ln k_{it})^2$	-0.0807***	-0.0811***	-0.0814***	-0.0811***	-0.179***	-0.182***	-0.179***	-0.183***
	(0.0103)	(0.0103)	(0.0117)	(0.0117)	(0.00960)	(0.00954)	(0.0101)	(0.00974)
(8) $\ln a_{it}^{\mathrm{INV}}$	0.115**	0.116***	0.0143	-0.00638	0.205***	0.222***	0.0993	0.0782
	(0.0520)	(0.0438)	(0.0491)	(0.0379)	(0.0678)	(0.0526)	(0.0782)	(0.0697)
(9) $\ln a_{it}^{\mathrm{R\&D}}$	0.0293***	0.0233***	0.0250***	0.0226***	0.0413***	0.0328***	0.0375***	0.0332***
	(0.00763)	(0.00597)	(0.00642)	(0.00560)	(0.0121)	(0.0103)	(0.0114)	(0.0109)
(10) $\ln a_{it}^{\mathrm{MNF}_e}$	0.0270***	0.0195***	0.0465***	0.0434***	0.00594	-0.0111	0.0289**	0.0224
	(0.00528)	(0.00606)	(0.00577)	(0.00544)	(0.0112)	(0.0102)	(0.0130)	(0.0144)
(11) $\ln a_{it}^{\mathrm{MNF}_o}$	0.00263	0.00534	-0.000942	0.00519	-0.00963	-0.00549	-0.0140	-0.00361
	(0.0109)	(0.0113)	(0.0103)	(0.00974)	(0.0109)	(0.0106)	(0.0114)	(0.0112)
(12) $\ln a_{it}^{\mathrm{POP}}$	-0.634	-0.885*	-1.141**	-1.110**	0.329	0.0790	-0.146	0.0460
	(0.466)	(0.502)	(0.508)	(0.477)	(0.361)	(0.422)	(0.434)	(0.433)
(13) τ_1	0.180***	0.127***	0.160***	0.134***	0.191***	0.128***	0.173***	0.146***
	(0.0205)	(0.0194)	(0.0197)	(0.0173)	(0.0344)	(0.0385)	(0.0340)	(0.0352)
(14) Hansen J p-val.	0.762	0.777	0.797	0.741	0.161	0.369	0.107	0.347
(15) 1st stage F	209.4	209.4	320.1	664.6	196.2	152.8	217.5	387.6
(16) #Obs.	99,996	99,996	99,996	98,880	99,996	99,996	99,996	98,880

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Columns 1–4 and 5–8 in Table 8.1 indicate the IV results for (4.1) under citation- and novelty-adjusted measures, respectively, with these additional controls on the RHS. The IVs are constructed using all the third, fourth, and fifth indirect collaborators because similar results are obtained when only one of them is used.[48]

[48] The first-stage F values (row 15) are reasonably large for all cases, indicating the strong relevance of the IVs. Hansen (1982)'s J-test indicates no evidence against the exogeneity of the IVs (row 14).

We find that the coefficient of $\ln k_{it}^{D}$ in (4.1) is significantly different between the baseline and the current specifications (under both the citation- and novelty-adjusted productivity) when the research scope at the firm/establishment and/or establishment size are controlled.[49] More specifically, up to around 30% of the estimated coefficients for $\ln k_{it}^{D}$ under the baseline specifications are accounted for by the firm/establishment-specific factors. Nevertheless, the signs and significance of the effect of $\ln k_{it}^{D}$ found in the baseline model persist.

In particular, the coefficients of research scope at both firm and establishment levels are positive and significant, which may relate to the effects of time-varying R&D resources at the firm and establishment levels that cannot be controlled by the local concentration of R&D expenditure $\ln a_{it}^{\mathrm{R\&D}}$.

Firm size $\ln f_{it}$ has negative significant effects on inventor productivity after controlling for the research scope $\ln s_{it}^{f}$ of the firm (columns 2 and 6). This may reflect the fact that larger firms have a comparably larger number of less skilled/experienced inventors and that they are necessarily assigned to some patent projects of the firm, resulting in the lower average productivity of an inventor.

At the establishment level, the size effect $\ln e_{it}$ and local concentration of inventors $\ln a_{it}^{\mathrm{INV}}$ become insignificant once the research scope of an establishment $\ln s_{it}^{e}$ is controlled (columns 3, 4, 7, and 8). Thus, the scale effect of an establishment is positive and is largely represented by the research scope of the establishment.

Next, columns 1–4 and 5–8 in Table 8.2 report the IV results for (4.6) under citation- and novelty-adjusted measures, respectively, with the additional controls on the RHS. The IVs are constructed by using all the third, fourth, and fifth indirect collaborators because similar results are obtained when only one of them is used.[50]

As in the case of (4.1) in Table 8.1, the research scope at both firm and establishment level has positive significant effects on the amount of differentiated knowledge of collaborators. The effects of both firm and establishment sizes are negative unlike the case of (4.1). A similar explanation as above applies here. Namely, a larger firm/establishment has a comparably larger number of less skilled/experienced inventors. Because most collaboration occurs within a firm/establishment, they are necessarily assigned to some patent projects of the firm/establishment, resulting in lower average differentiated knowledge of an inventor. The positive effect of the research scope of the firm/establishment (possibly representing the R&D resource) constitutes up to 8–9% of the estimated coefficients of $\ln \Delta n_{it}$ in the baseline case. Nevertheless, the signs and significance of the effect of $\ln \Delta n_{it}$ found in the baseline model persist.

[49]This is based on the Wald test in the generalized method of moments estimation, which simultaneously estimates the baseline and current models with the 2SLS weighting matrix.

[50]The first-stage F values (row 15) are reasonably large, indicating the strong relevance of the IVs. Hansen (1982)'s J-test indicates no evidence against the exogeneity of the IVs (row 14).

Table 8.2: Regression results for (4.6) with firm- and establishment-specific factors

	Citations				Novelty			
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) $\ln \Delta n_{it}$	1.407***	1.390***	1.292***	1.257***	1.775***	1.737***	1.633***	1.567***
	(0.0864)	(0.0847)	(0.0842)	(0.0667)	(0.113)	(0.112)	(0.115)	(0.0927)
(2) $\ln f_{it}$	0.0111	-0.236			0.0493	-0.517***		
	(0.100)	(0.144)			(0.160)	(0.199)		
(3) $\ln s^f_{it}$		0.210***				0.481***		
		(0.0642)				(0.0877)		
(4) $\ln e_{it}$			0.195***	-0.0952			0.256***	-0.360**
			(0.0387)	(0.124)			(0.0655)	(0.159)
(5) $\ln s^e_{it}$				0.175**				0.377***
				(0.0718)				(0.0767)
(6) $\ln k_{it}$	-0.0215	-0.0268	-0.0153	-0.0237	-0.0323	-0.0445	-0.0248	-0.0411
	(0.0744)	(0.0700)	(0.0741)	(0.0680)	(0.0722)	(0.0633)	(0.0714)	(0.0591)
(7) $(\ln k_{it})^2$	0.237***	0.235***	0.214***	0.210***	0.280***	0.275***	0.252***	0.245***
	(0.0213)	(0.0206)	(0.0214)	(0.0231)	(0.0240)	(0.0226)	(0.0255)	(0.0274)
(8) $\ln a^{\mathrm{INV}}_{it}$	0.0187	0.00671	-0.0568	-0.0390	0.0820	0.0543	-0.0103	0.00241
	(0.0410)	(0.0422)	(0.0347)	(0.0400)	(0.0630)	(0.0496)	(0.0655)	(0.0677)
(9) $\ln a^{\mathrm{R\&D}}_{it}$	-0.000660	-0.00895	-0.00413	-0.00673	0.0154	-0.00356	0.0115	0.00359
	(0.00682)	(0.00723)	(0.00665)	(0.00690)	(0.0113)	(0.0123)	(0.0107)	(0.00999)
(10) $\ln a^{\mathrm{MNF}_e}_{it}$	-0.00294	-0.00938	0.0157	0.0193	-0.0319	-0.0466**	-0.00629	-0.00618
	(0.0149)	(0.0151)	(0.0154)	(0.0147)	(0.0219)	(0.0205)	(0.0249)	(0.0278)
(11) $\ln a^{\mathrm{MNF}_o}_{it}$	0.00189	0.00495	-0.00133	0.00889	-0.0167**	-0.00971	-0.0206***	0.000505
	(0.0111)	(0.0114)	(0.0101)	(0.00849)	(0.00832)	(0.00960)	(0.00772)	(0.00669)
(12) $\ln a^{\mathrm{POP}}_{it}$	-0.921	-1.257	-1.307	-1.148	0.985	0.215	0.519	0.781
	(1.334)	(1.243)	(1.313)	(1.183)	(1.365)	(1.149)	(1.250)	(1.036)
(13) τ_1	0.524***	0.438***	0.480***	0.419***	0.826***	0.629***	0.775***	0.638***
	(0.0619)	(0.0425)	(0.0566)	(0.0418)	(0.0512)	(0.0276)	(0.0381)	(0.0339)
(14) Hansen J p-val.	0.325	0.323	0.310	0.305	0.451	0.468	0.477	0.504
(15) 1st stage F	198.1	193.5	188.6	175.5	198.1	193.5	188.6	175.5
(16) #Obs.	82,048	82,048	82,048	81,270	82,048	82,048	82,048	81,270

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

8.2 Research affiliations of firms and establishments

Although more than 80% of collaboration occurs within a firm and within an establishment on average, the entire collaboration often involves multiple firms and establishments. The productivity of an inventor and the amount of differentiated knowledge of collaborators may be influenced by the time-varying factors specific to the set of firms and establishments that affiliate for joint R&D in which they are involved. For example, variation in productivity as well as the differentiated knowledge of potential collaborators in the affiliated firms may substantially depend on the capability of the chief executive officer of the leading firm, the presence or absence of star inventors in the affiliation, the research funds allocated to the affiliation in a given period, and so on. These factors specific to affiliated firms and establishments may also be correlated with the magnitude of collaborator recombination of an inventor. For example, newly available R&D resources to a given affiliation of firms or establishments may induce an unusual recombination of collaborators to improve complementarity within new research projects. Consequently, the omission of these factors may result in a bias of the estimated coefficient of $\ln k_{it}^D$ in (4.1) and that of $\ln \Delta n_{it}$ in (4.6).

To evaluate the influence of such factors in models (4.1) and (4.6), we consider random counterfactual choices of collaborators for each inventor conditional on the actual number as well as firm/establishment affiliation of each of their collaborators.

Suppose that among the n_{it} collaborators of inventor i in period t, n_{it}^A belongs to firm A, n_{it}^B belongs to firm B, and so on, where $n_{it} = \sum_j n_{it}^j$. Then, these n_{it} collaborators are replaced by n_{it}^A randomly chosen collaborators without replacement (according to the uniform probability distribution) from firm A, n_{it}^B from firm B, and so on, respectively; however, the second or closer indirect and direct collaborators of i are excluded from the selection to mitigate the reflection problem. Alternatively, the counterfactual collaborators of inventor i may be chosen conditional on the actual establishment (rather than firm) affiliation of each collaborator of i.

We construct 1,000 sets of counterfactual collaboration patterns in this way and compute the counterfactual k_{it}^D under each. Models (4.1) and (4.6) are estimated by OLS under each counterfactual value of k_{it}^D. Our interest is the extent to which the causality from $\ln k_{it}^D$ to $\ln y_{it}$ in (4.1) and from $\ln \Delta n_{it}$ to $\ln k_{it}^D$ can be attributed to factors specific to the affiliated firms and establishments rather than the direct effects of knowledge exchange and the collaborator recombination, respectively.[51]

Table 8.3 summarizes the estimated values of the coefficient β of $\ln k_{it}^D$ of model (4.1) and that of the coefficient $\tilde{\beta}$ of $\ln \Delta n_{it}$ under counterfactual data and compares them with their IV estimates under the actual data.[52]

As for model (4.1), the affiliated firm-specific factors are significant, on average, under the citation-adjusted measures (columns 1 and 2); however, their magnitude is less than 10% of the IV estimate of β under the actual data (column 2 of tables 7.1 and 8.1). Under the novelty-adjusted measures (columns 3 and 4), the affiliated firm-specific factors are at most weakly significant on average (without firm-specific controls) though they account for, on average, 3.6% of the IV estimate of β under the actual data.

The affiliated establishment-specific factors are significant, on average, under both quality measures (columns 5–8) and account for 17.7% and 15.9% of the total effect of $\ln k_{it}^D$ under the citation-adjusted measure without and with establishment-specific controls, respectively. The corresponding numbers under the novelty-adjusted measure are 8.6% and 6.3%, respectively. Because citations are more likely among affiliated firms and establishments, both affiliated firm- and establishment-specific factors matter

[51]There may be a concern for endogeneity if there were positive assortative matching between an inventor and the affiliated firms and establishments. In that case, the estimation bias for the coefficients of $\ln k_{it}^D$ and $\ln \Delta n_{it}$ is expected to be upward, which makes our robustness check conservative.

[52]Here, we consider the IV estimates based on the third to fifth indirect collaborators.

more for citation-adjusted productivity.

Nonetheless, more than 80% of the original estimates of β and $\tilde{\beta}$ still remain after removing the affiliated firm/establishment-specific factors. We find this to be strong supportive evidence for the BF mechanism of knowledge creation.

Table 8.3: Effects of affiliated firm/establishment-specific factors

		Firms				Establishments			
		Citations		Novelty		Citations		Novelty	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model (4.1)	Average estimate of β	0.028***	0.026***	0.012*	0.008	0.051***	0.032***	0.029***	0.015*
	(S.D.)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
	Share in the actual	0.099	0.094	0.036	0.029	0.177	0.159	0.086	0.063
Model (4.6)	Average estimate of $\tilde{\beta}$	-0.028	-0.033**	0.183***	0.175***	0.052***	0.012	0.267***	0.211***
	(S.D.)	(0.010)	(0.010)	(0.017)	(0.017)	(0.009)	(0.009)	(0.014)	(0.009)
	Share in the actual	-0.021	-0.025	0.106	0.101	0.038	0.010	0.155	0.134
Additional controls	$\ln f_{it}$, $\ln s_{it}^{f}$		✓		✓				
	$\ln e_{it}$, $\ln s_{it}^{e}$						✓		✓

Columns 1–4 and 5–8 report the results conditional on the actual firm- and establishment-level affiliations, respectively. The productivity measures are quality-adjusted in columns 1, 2, 5, and 6 and novelty-adjusted in columns 3, 4, 7, and 8. Columns of odd numbers show the results under (4.1), with the actual k_{it}^{D} replaced by the counterfactual one. The columns of even numbers show the results with additional controls of the size and technological scope of a firm or establishment. The "share in the actual" indicates the ratio of the average estimate of β under the counterfactual data to the estimate of that under the actual data. The estimated values of β under the actual data are the IV estimates, with IVs constructed by using the third through fifth indirect collaborators. Specifically, these are 0.286 for columns 1 and 5, 0.344 for columns 3 and 7 (refer to Table 7.1), and 0.271, 0.276, 0.201, and 0.229 for columns 2, 6, 4, and 8, respectively (refer to Table 8.1). The superscripts $***$, $**$, and $*$ indicate that the estimated values of β and $\tilde{\beta}$ are significant at the 0.01, 0.05, and 0.1 levels, respectively, on average.

8.3 Other robustness analyses

Finally, this section briefly discusses the results of other robustness analyses. In all cases, the signs and significance of the estimated coefficients of $\ln k_{it}^{D}$ in (4.1) and $\ln \Delta n_{it}$ in (4.6) are consistent with our baseline results shown in tables 7.1 and 7.3.

Alternative productivity measures – The regressions for (4.1) and (4.6) are conducted under four alternative measures of inventor productivity, where the output g_j of patent j in (2.1) is given by (*i*) cited count in five years from publication, (*ii*) technological novelty based on the IPC subclass, and (*iii*) count of patent claims,[53] or (*iv*) count of patents, i.e., $g_j = 1$ for all j. See Appendix E.1 for the details.

Differentiated knowledge of collaborators by IPC – The differentiated knowledge of collaborators may be defined by the IPC categories as $k_{it}^{D} = \frac{1}{n_{it}} \sum_{j \in N_{it}} \left| S_{jt} \backslash S_{it} \right|$ instead of the productivity-based measures; however, it corresponds less precisely to the knowledge creation function (3.1). See Appendix E.2 for the details.[54]

[53]Each claim indicates an aspect of the patent to be protected; thus, its count reflects the technological novelty *within a patent*. Although the claims are made by applicants, this is not an entirely subjective measure of quality because each claim incurs monetary costs.

[54]The qualitative results remain the same if the IPC subclass instead of subgroup is adopted to define the differentiated knowledge of collaborators as well as the cumulative research scope of inventors.

Alternative radius values for local concentration – We consider alternative radius values (5, 10, and 20 km) to quantify the magnitude of local concentration of inventors, R&D expenditure, manufacturing, and population around each individual inventor. The direct effects of local concentrations on the outcome variables are generally robust under the alternative radius values, with some exceptions. Specifically, for (4.1), the effects of local concentrations are spatially confined for the inventor, R&D expenditure, and manufacturing concentrations in the sense that the effect is significant up to a 5 km radius (if significant at all). The effect of residential population concentration $\ln a_{it}^{\mathrm{POP}}$ is insignificant for all radius values 5–20 km. For (4.6), the negative effects of $\ln a_{it}^{\mathrm{INV}}$ and $\ln a_{it}^{\mathrm{MNF}_e}$ persist for larger radius values, possibly reflecting the tougher competition for human resources with co-localizing R&D activities as well as the manufacturing sector. See Appendix E.3 for the details.

IVs based on indirect collaborators in different firms – Finally, we consider alternative IVs for $\ln k_{it}^{D}$ and $\ln \Delta n_{it}$, which are the same as $\ln k_{it}^{D,\mathrm{IV}_\ell}$ in (6.3) and $\ln \Delta n_{it}^{\mathrm{IV}_\ell}$ in (6.4) except that the inventors in the same firm as i are excluded from N_{it}^{ℓ}. Using these IVs, we can mitigate the influence of unobserved firm-specific factors associated with the IVs that may correlate with the error term.

The results remain qualitatively the same as those from the baseline analyses shown in tables 7.1 and 7.3 except that the estimated coefficient for $\ln k_{it}^{D}$ in (4.1) is insignificant under the IVs constructed from the fourth and fifth indirect collaborators for the case of citation-adjusted productivity. This is not surprising provided that the correlations of worker productivity among firms attenuate rather quickly as firms become far from each other on the investment network as discussed in Section 6.3.

Nevertheless, weak IVs are not found under the novelty-adjusted productivity measures. Moreover, for all cases, if we use both baseline and present IVs constructed from the ℓ-th indirect collaborators for $\ell = 3,4$ and 5, then the null hypothesis of the Hansen (1982)'s J-test is not rejected, which suggests that unobserved firm-specific factors are of minor concern. For (4.6), we have qualitatively the same results under the alternative IVs as those obtained in Section 7.3. See Appendix E.4 for the details.

9 Discussion and further research directions

We have shown evidence consistent with the polyadic collaborative knowledge creation mechanism proposed by Berliant and Fujita (2008). To our knowledge, our work is the first attempt to provide micro-econometric evidence for knowledge creation at the individual inventor level under endogenous collaboration.

We have also addressed two major counterfactual aspects of the BF model, guided by Horii's (2012) result. One is that each inventor in their model belongs to a fixed network component in a typical steady state, meaning that polyadic interactions happen only within a given set of collaborators. However, in the data, the set of collaborators evolves for each agent over time and the intertemporal recombination of collaborators is found to revise inventors' technological expertise by meeting new agents and adopting their differentiated knowledge.

The other is that inventors in their model face no imitation or obsolescence of their technological knowledge because the potential knowledge is infinite and symmetric. In reality, however, the past research experience of an inventor may hinder their future invention (except for superstars with particularly high research profiles) as old knowledge quickly reduces to obsolescence and is subject to imitation. Thus, if inventors stick to their past achievements, then they most likely lose their present level of creativity in the long run. If instead agents are willing to explore new research directions by meeting new collaborators with different backgrounds from theirs, then they are more likely to keep their creativity by shifting their technological expertise to unexplored niches. We have explained this causal relation by estimating the second and third regression models, (4.5) and (4.6), in addition to the original BF model (4.1). Specifically, collaborator recombination is found to be an effective means for raising the quality of collaborators' differentiated knowledge and thereby enhancing the quality of the inventor's output.

This evidence has important policy implications. For example, firms, cities, regions, and countries that promote encounters and collaboration among individual inventors across organizations and institutions, despite the possibility of imitations and undesired diffusions, may have better chances to foster innovation there. While lower organizational and institutional barriers for research collaboration are not incompatible with the protection of intellectual property by patents, our finding supports more active coordination than divisions among researchers to encourage innovation.[55]

Among a number of short- and long-run extensions, we touch on three. First, it is an obvious interest to further investigate the roles of firms and establishments in R&D activities. Because the financial resources for R&D are typically provided by firms, firm-specific patterns of collaborations and R&D policies could affect the productivity of individual inventors and firms.[56] By matching the addresses of establishments in the patent database with those of the Census of Manufacturers, it is also possible to investigate the impact of patent development on firm productivity.

Second, non-technological diversity among collaborators in terms of, for example, gender, age, and cultural background, may affect productivity. For example, Østergaard et al. (2011) and Inui et al. (2014) found a positive influence of gender diversity

[55]See Boldrin and Levine (2013) for a related survey of the literature arguing that the patent system hinders rather than promotes innovation.

[56]See Akcigit and Kerr (2018) for an initial attempt in this direction as they distinguish between R&D that is internal and external to firms and study the firm dynamics that arise from this distinction.

on innovation productivity of Danish and Japanese firms, respectively.

Finally, it is intriguing to explore the differences in the location patterns of R&D activities and industries. It is argued that large cities are typically associated with a concentration of knowledge-intensive activities (e.g., Davis and Dingel, 2019). However, the fundamental distinction between knowledge-intensive and non-intensive activities has not been made clear thus far. From our findings, knowledge-intensive R&D activities are clearly expected to be more concentrated geographically given their incentive for frequent collaborator recombination than industrial activities whose concentrations are typically induced by input–output linkages, demand, and production externalities.

Figure 9.1(a) plots the aggregate novelty-adjusted patent output and manufacturing output against the population size of a UA in period 1, where all values are expressed by shares in all UAs.[57] The solid and dashed lines indicate the fitted OLS lines for the patent and manufacturing output plots, respectively. While both plots indicate increasing per capita productivity in the UA size, it is substantially more so for patent output: Doubling the population size of a UA raises R&D productivity by 2.5 times[58] while raising manufacturing productivity only by 1.2 times.

Figure 9.1(b) plots the diversity in IPC subclasses of applied patents as well as industrial diversity in terms of the number of four-digit Japanese SIC manufacturing industries against the population size of UAs in 2000, where all values are in shares again.[59] Comparing UAs in terms of the diversity in IPC subclass and SIC four-digit industry categories is reasonable because they are comparable in terms of the total number of active categories, which is 608 for the former and 562 for the latter. The solid and dashed lines indicate the fitted OLS lines for the patent class and industrial diversity plots, respectively. While diversity is increasing in the population size of a UA for both patent categories and manufacturing industries, the increase of the former is substantially more. Thus, doubling the population size of a UA almost doubles the diversity in the technological category of patents applied in the UA, whereas it only increases the industrial diversity by 55%. Thus, while a larger UA is associated with both larger intensive (i.e., per capita output) and extensive margins (i.e., diversity) in both R&D and production activities, this tendency is substantially stronger for the former.

These findings suggest a positive association between population concentration and matching externalities that promote collaborator recombination in large cities.[60]

[57] The location of a patent is identified by that of the patent applicant. Manufacturing output is obtained from the micro data of the Census of Manufacturers in 2000.

[58] Estimated elasticities of patent output with respect to UA population are similar among alternative output measures: under IPC subclass and cited count, they are 1.516 and 1.458, respectively,

[59] The industrial diversity of a given UA is defined as the number of four-digit manufacturing industries that have positive employment in the UA.

[60] See Agrawal et al. (2017); Perlman (2016) for evidence of the effect of transport costs on R&D agglomeration.

However, the mechanism behind the difference between R&D and industrial location patterns has not been fully explored either theoretically or empirically, and this remains a future research subject.[61]

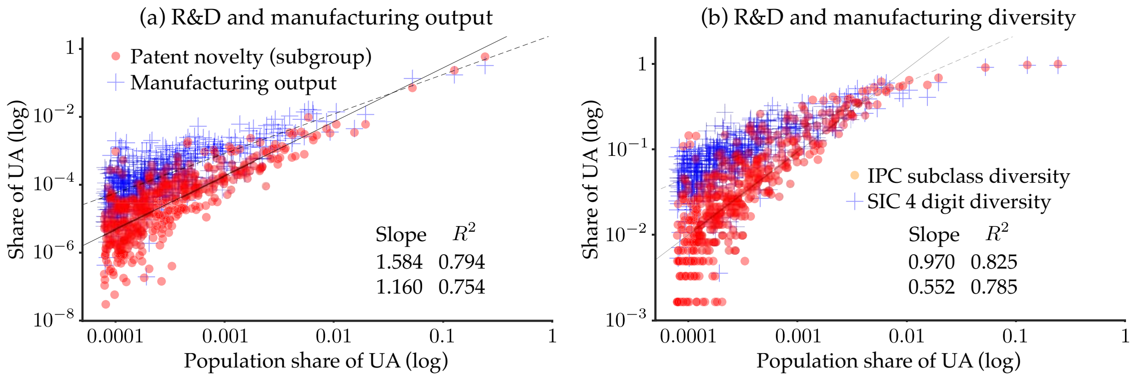


Figure 9.1: Industrial and research outputs and diversities in UAs in 2000

[61] It is also possible to identify the geographic scope of collaborations, e.g., within an establishment, a district, a metropolitan area, and so on. See Gordon (2013) for an initial attempt in this direction.

References

Abrams, David S., Ufuk Akcigit, and Jillian Popadak, "Patent value and citations: Creative destruction or strategic disruption?," 2013. Discussion paper No. 19647, National Bureau of Economic Research.

Aghion, Philippe and Peter Howitt, "A model of growth through creative destruction," *Econometrica*, 1992, *60* (2), 323–351.

Agrawal, Ajay, Alberto Galasso, and Alexander Qettl, "Roads and innovation," *Review of Economics and Statistics*, 2017, *99* (3), 417–434.

Akcigit, Ufuk and William R. Kerr, "Growth through heterogeneous innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

_ , John Grigsby, and Tom Nicholas, "The rise of American ingenuity: Innovation and inventors of the Golden Age," 2017. Discussion paper No. 23047, National Bureau of Economic Research.

_ , Santiago Caicedo, Ernest Miguelez, Stefanie Santcheva, and Valerio Sterzi, "Dancing with the stars: Innovation through interactions," 2018. Discussion paper No. 24466, National Bureau of Economic Research.

Artificial Life Laboratory, Inc., "Patent database for research," 2018.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang, "Superstar extinction," *The Quarterly Journal of Economics*, 2010, *125* (2), 549–589.

Bartolucci, Cristian and Francesco Devicienti, "Better workers move to better firms: A simple test to identify sorting," 2013. Discussion paper No. 7601, Institute of Labor Economics (IZA).

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud, "Productive cities: Sorting, selection, and agglomeration," *Journal of Political Economy*, 2014, *122* (3), 507–553.

Berliant, Marcus and Masahisa Fujita, "Knowledge creation as a square danc on the Hilbert cube," *International Economic Review*, 2008, *49* (4), 1251–1295.

Bettencourt, Luís M. A., José Lobo, Dirk Helbing, Cristian Kühnert, and Geoffrey B. West, "Growth, innovation, scaling, and the pace of life in cities," *Proceedings of the National Academy of Sciences of the United States of America*, 2007, *104* (7), 7301–7306.

Boldrin, Michele and David K. Levine, "The case against patents," *Journal of Economic Perspectives*, 2013, *27* (1), 3–22.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin, "Identification of peer effects through social networks," *Journal of Econometrics*, 2009, *150* (1), 41–55.

Breschi, Stefano, Francesco Lissoni, and Franco Malerba, "Knowledge-relatedness in firm technological diversification," *Research Policy*, 2003, *32* (1), 69–87.

Cai, Jie and Nan Li, "Growth through inter-sectoral knowledge linkages," *Review of Economic Studies*, 2019, *forthcoming.*

Calvó-Armengol, Antoni, Elenora Patacchini, and Yves Zenou, "Peer effects and social network in education," *Review of Economic Studies*, 2009, *76* (4), 1239–1267.

Chang, Howard F., "Patent scope, antitrust policy, and cumulative innovation," *RAND Journal of Economics*, 1995, *26* (1), 34–57.

Chu, Angus C., "Effects of blocking patents on R&D: A quantitative DGE analysis," *Journal of Economi Growth*, 2009, *14* (1), 55–78.

Coe, Davit T. and Elhanan Helpman, "International R&D spillovers," *European Economic Review*, 1995, *39* (5), 859–887.

Comola, Margherita and Silvia Prina, "Treatment effects accounting for network changes," 2017. Unpublished manuscript. Paris School of Economics.

Cozzi, Guido and Silvia Galli, "Sequential R&D and blocking patents in the dynamics of growth," *Journal of Economic Growth*, 2014, *19* (2), 183–219.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum, "Matching in cities," 2018. Discussion paper No.25227, National Bureau of Economic Research.

Davis, Donald R. and Jonathan I. Dingel, "A spatial knowledge economy," *American Economic Review*, 2019, *109* (1), 153–170.

Duranton, Gilles and Diego Puga, "Nursery cities: Urban diversity, process innovation, and the life cycle of products," *American Economic Review*, 2001, *91* (5), 1454–1477.

Eeckhout, Jan and Philipp Kircher, "Assortative matching with large firms," *Econometrica*, 2018, *86* (1), 85–132.

Garcia-Vega, Maria, "Does technological diversification promote innovation? An empirical analysis for European firms," *Research Policy*, 2006, *35* (2), 230–246.

Gaubert, Cecile, "Firm sorting and agglomeration," *American Economic Review*, 2018, *108* (11), 3117–3153.

Giorgi, Giacomo De, Michele Pellizza, and Silvia Redaelli, "Identification of social interactions through partially overlapping peer groups," *American Economic Journal: Applied Economics*, 2010, *2* (2), 241–275.

Glass, Amy Jocelyn and Kamal Saggi, "Licensing versus direct investment: Implications for economic growth," *Journal of International Economics*, 2002, *56* (1), 131–153.

Goldsmith-Pinkham, Paul and Guido W. Imbens, "Social networks and identification of peer effects," *Journal of Business & Economic Statistics*, 2013, *31* (3), 253–264.

Gordon, Peter, "Thinking about economic growth: Cities, networs, creativity and supply chains for ideas," *Annals of Regional Science*, 2013, *50* (3), 667–684.

Grossman, Gene M. and Carl Shapiro, "Dynanic R&D competition," *The Economic Journal*, 1978, *97* (386), 372–387.

_ **and Elhanan Helpman**, *Innovation and Growth in the Global Economy*, Cambridge, MA: The MIT Press, 1991.

Hansen, Lars Peter, "Large sample properties of generalized method of moments estimators," *Econometrica*, 1982, *50* (4), 1029–1054.

Horii, Ryo, "Wants and past knowledge: Growth cycles iwth emerging industries," *Journal of Economic Dynamics and Control*, 2012, *36* (2), 220–238.

Hsieh, Chih-Sheng and Lung-Fei Lee, "A social interactions model with endogenous friendship formation," *Journal of Applied Econometrics*, 2016, *31* (2), 1–21.

Huo, Dong and Kazuyuki Motohashi, "Understanding two types of technological diversity and their effects on the technological value of outcomes from bilateral inter-firm R&D alliances," 2015. Discussion paper No. 15-E-06, The Research Institute of Economy, Trade and Industry (RIETI).

Iaria, Alessandro, Carlo Schwartz, and Fabian Waldinger, "Frontier knowledge and scientific production: Evidence from the collapse of international science," *The Quarterly Journal of Economics*, 2018, *133* (2), 927–991.

Inoue, Hiroyasu, Kentaro Nakajima, and Yukiko Saito Umeno, "Innovation and collaboration patterns between research establishments," 2015. Discussion paper No. 15-E-049, The Research Institute of Economy, Trade and Industry (RIETI).

_ , _ , and Yukiko Umeno Saito, "Localization of knowledge-creating establishments," *Japan and the World Economy*, 2017, *43*, 23–29.

Inui, Tomohiko, Makiko Nakamuro, Kazuma Edamura, and Junko Ozawa, "Impact of diversity and work-life balance (in Japanese)," 2014. Discussion paper No. 14-J-055, Research Institute of Economy, Trade and Industry (RIETI).

Jackson, Matthew O., Brian W. Rogers, and Yves Zenou, "The economic consequence of social-network structure," *Journal of Economic Literature*, 2017, *55* (1), 49–95.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, "Geographic localization of knowledge spillovers as evidenced by patent citations," *The Quarterly Journal of Economics*, 1993, *108* (3), 577–598.

Johnsson, Ida and Hyungsik Roger Moon, "Estimation of peer effects in endogenous social networks: Control function approach," 2019. arXiv:econ.EM/1709.10024v2.

Jovanovic, Boyan and Rafael Rob, "The diffusion and growth of knowledge," *Review of Economic Studies*, 1989, *56* (4), 569–582.

Kerr, William R. and Scott Duke Kominers, "Agglomerative forces and cluster shapes," *Review of Economics and Statistics*, 2015, *97* (4), 877–899.

Klette, Tor Jakob and Samuel Kortum, "Innovating firms and aggregate innovation," *Journal of Political Economy*, 2004, *112* (5), 986–1018.

König, Michael D., Xiaodong Liu, and Yves Zenou, "R&D networks: Theory, empirics and policy implications," *Review of Economics and Statistics*, 2019, *forthcoming*.

Kortum, Samuel, "Research, patenting, and technological change," *Econometrica*, 1997, *65* (6), 1389–1419.

Lentz, Rasmus and Dale T. Mortensen, "An empirical model of growth through product innovation," *Econometrica*, 2008, *76* (6), 1317–1373.

Li, Tong and Li Zhao, "A partial identification subnetwork approach to discrete games in large networks: An application to quantifying peer effects," 2016. Unpublished manuscript, Vanderbilt University.

Lucas, Robert E. and Benjamin Moll, "Knowledge growth and the allocation of time," *Journal of Political Economy*, 2014, *122* (1), 1–51.

Manski, Charles F., "Identification of endogenous social effects: The reflection problem," *Review of Economic Studies*, 1993, *60* (3), 531–542.

Matutes, Carmen, Pierre Regibeau, and Katharine Rockett, "Optimal patent design and the diffusion of innovations," *RAND Journal of Economics*, 1996, *27* (1), 60–83.

Mendes, Rute, Gerard J. van den Berg, and Maarten Lindeboom, "An empirical assessment of assortative matching in the labor market," *Labour Economics*, 2010, *17* (6), 919–929.

Ministry of Economy, Trade and Industry of Japan, *Census of Manufactures* 1995, 2000, 2005.

_ , *METI Basic Survey of Japanese Business Structure and Activities* 1995-2010.

Ministry of Internal Affairs and Communications of Japan, *Population Census (Tabulation for standard area mesh)* 1995, 2000, 2005.

_ , *Establishment and Enterprise Census* 1996, 2001, 2006.

_ , *Survey of Research and Development* 1997-2010.

_ , *Economic Census for Business Frame* 2009.

_ , *Population Census (Tabulation for standard area mesh)* 2010.

Mori, Tomoya and Alessandro Turrini, "Skills, agglomeration and segmentation," *European Economic Review*, 2005, *49* (1), 201–225.

Murata, Yasusada, Ryosuke Okamoto Ryo Nakajima, and Ryuichi Tamura, "Localized knowledge spillovers and patent citations: A distanced-based approach," *Review of Economics and Statistics*, 2014, *96* (5), 967–985.

Nakajima, Kentaro, Yukiko Umeno Saito, and Iichiro Uesugi, "Localization of interfirm transaction relationships and industry agglomeration," April 2012. Discussion paper 12-E-023, Research Institute of Economy, Trade and Industry.

Olsson, Ola, "Knowledge as a set in idea space: An epistemological view on growth," *Journal of Economic Growth*, 2000, *5* (3), 253–275.

_ , "Technological opportunity and growth," *Journal of Economic Growth*, 2005, *10* (1), 31–53.

Østergaard, Christian R., Bram Timmermans, and Kari Kristinsson, "Does a different view create something new? The effect of employee diversity on innovation," *Research Policy*, 2011, *40* (3), 500–509.

Otazawa, Toshimori, Yuki Ohira, and Jos van Ommeren, "Inter-firm transaction networks and location in a city," August 2018. Discusssion paper 18-E-054, Research Institute of Economy, Trade and Industry.

Panebianco, Fabrizio, Thierry Verdier, and Yves Zenou, "Innovation, pricing and targeting in networks," 2016. Discussion paper No.11398, Center for Economic Policy Research.

Patacchini, Eleonora, Edoardo Rainone, and Yves Zenou, "Heterogeneous peer effects in education," *Journal of Economic Behavior & Organization*, 2017, *134* (C), 190–227.

Perla, Jesse and Christopher Tonetti, "Equilibrium imitation and growth," *Journal of Political Economy*, 2018, *122* (1), 52–76.

Perlman, Elizabeth Ruth, "Dense enough to be brilliant: Patents, urbanization, and transportation in nineteenth century America market access, and information flows," 2016. Unpublished manuscript, Boston University.

Romer, Paul M., "Endogenous technological change," *Journal of Political Economy*, 1990, *98* (5), 71–102.

Scherer, Frederic M., "Inter-industry technology flows and productivity growth," *Review of Economics and Statistics*, 1982, *64* (4), 627–634.

Schotchmer, Suzanne, "Protecting early inventors: Should second-generation products be patentable?," *RAND Journal of Economics*, 1996, *27* (2), 322–331.

Shell, Karl, "Toward a theory of inventive activitgy and capital accumulation," *American Economic Review*, 1966, *56* (1/2), 62–68.

Tanaka, Hitoshi, Tatsuro Iwaisako, and Koichi Futagami, "Dynamic analysis of innovation and international transfer of technology through licensing," *Journal of International Economics*, September 2007, *73* (1), 189–212.

Thompson, Peter and Melanie Fox-Kean, "Patent citations and the geography of knowledge spillovers: A reassessment," *The American Economic Review*, 2005, *95* (1), 450–460.

Tokyo Shoko Research, "Kigyo Zaimu Joho and Sokan File," 2014.

Trajitenberg, Manuel, "A penny for your quotes: Patent citations and the value of innovations," in Adam B. Jaffe and Manuel Trajtenberg, eds., *Patents, Citations & Innovations*, Cambridge, MA: The MIT Press, 2002, chapter 2, pp. 27–49.

Waldinger, Fabian, "Quality matters: The expulsion of professors and the consequences for PhD student outcomes in Nazi Germany," *Journal of Political Economy*, 2010, *118* (4), 787–831.

_ , "Peer effects in science: Evidence from the Dismissal of scientists in Nazi Germany," *The Review of Economic Studies*, April 2012, *79* (2), 838–861.

Weitzman, Martin L., "Recombinant growth," *Quarterly Journal of Economics*, 1998, *113* (2), 331–360.

Yang, Guifang and Keith E. Maskus, "Intellectual property rights, licensing, and innovation in an endogenous product-cycle model," *Journal of International Economics*, 2001, *53* (1), 169–187.

APPENDIX

A Details of data

A.1 Descriptive statistics for the basic data

Table A.1: Descriptive statistics of basic variables

		Period	
		(1)	(2)
Variable		1	2
(1) Number of patents	$\left\lvert\bigcup_{i\in I}\mathcal{G}_{it}\right\rvert$	1,758,780	1,546,596
(2) Number of IPC classes		120	122
(3) Number of IPC subclasses		608	615
(4) Number of IPC subgroups	$\left\lvert\bigcup_{i\in I}S_{it}\right\rvert$	40,691	38,339
(5) Number of inventors in period *t*	$\lvert I_t\rvert$	1,208,197	1,094,789
(6) Number of inventors active in all periods	$\lvert I\rvert$	107,724	107,724
(7) Number of inventors per patent	$\lvert G_{jt}\rvert$	2.193 (1.538)	2.244 (1.609)
(8) Share of collaborating inventors	$\left\lvert\{i\in I_t : \lvert N_{it}\rvert > 0\}\right\rvert/\lvert I_t\rvert$	0.896	0.868
(9) Number of collaborators per inventor	$\lvert N_{it}\rvert$	8.518 (9.321)	6.323 (7.579)
(10) Number of new collaborators per inventor	Δn_{it}	6.893 (7.907)	4.354 (5.848)
(11) Number of patents per inventor	$\lvert\mathcal{G}_{it}\rvert$	10.66 (16.21)	6.858 (11.95)
(12) Number of IPC sections per inventor		1.812 (0.952)	1.533 (0.799)
(13) Number of IPC classes per inventor		2.473 (1.788)	1.918 (1.381)
(14) Number of IPC subclasses per inventor		2.984 (2.409)	2.241 (1.874)
(15) Number of IPC subgroups per inventor	$\lvert S_{it}\rvert$	5.471 (5.223)	3.713 (4.026)
(16) Size of cumulative IPC subgroups per inventor	$\lvert\cup_{t'<t}S_{it'}\rvert$	4.550 (4.659)	8.958 (7.582)

Numbers in rows 7–16 are the average values with the standard deviations in parentheses.

Table A.2: Descriptive statistics of productivity variables

Productivity measure		**Cited counts**		**Novelty**	
		(1)	(2)	(3)	(4)
Period		1	2	1	2
(1) Output of a patent	g_{jt}	1.535 (2.527)	1.423 (3.850)	0.013 (0.056)	0.009 (0.049)
(2) Productivity of an inventor	$\bar{y}_{it}$	7.906 (16.83)	5.048 (163.31)	0.047 (0.134)	0.024 (0.084)
(3) Pairwise productivity of an inventor	y_{it}	1.389 (3.160)	1.728 (175.04)	0.009 (0.049)	0.006 (0.032)
(4) Avg. diff. knowledge of collaborators	k^D_{it}	1.411 (7.520)	1.053 (4.539)	0.008 (0.043)	0.005 (0.034)

Numbers are the average values with standard deviations in parentheses.

A.2 IPC

The IPC classifies technologies into eight sections: A (human necessities), B (performing operations; transporting),…, H (electricity). These sections are divided into classes such as A01 (agriculture; forestry; animal husbandry; hunting; trapping; fishing) and then into subclasses such as A01C (planting; sowing; fertilizing). Each subclass is further divided into groups, e.g., A01C1 (apparatus, or methods of use thereof, for testing or treating seed, roots, or the like, prior to sowing or planting), and then into subgroups, e.g., A01C1/06 (coating or dressing seed) and A01C1/08 (immunizing seed). The IPC's labeling scheme is consistent over time, and a newly introduced category is basically associated with a new technology (e.g., the classes B81 for microtechnology and B82 for nanotechnology introduced in 2000). As another example, the shale revolution in the late 2000s in the United States was made possible by some key innovations in excavation technology that mainly belong to a new subclass C09K (compositions for drilling of boreholes or wells; compositions for treating boreholes or wells) that was split from E21B (earth or rock drilling; obtaining oil, gas, water, soluble or meltable materials or a slurry of minerals from wells) in 2006. If there are no fundamental changes in technology in a given category, the classification remains the same (e.g., A47C for furniture; domestic articles or appliances; coffee mills; spice mills; suction cleaners in general). Taken together, the set of technological categories specified in the IPC at a given point in time roughly represents the set of the state-of-the-art technologies at that time, and hence makes an appropriate proxy for the set of technological knowledge.

We have 121, 609, and 40,691 (123, 616, and 38,339) relevant IPC classes, subclasses, and subgroups, respectively for period 1 (period 2), associated with the applied patents in our data.

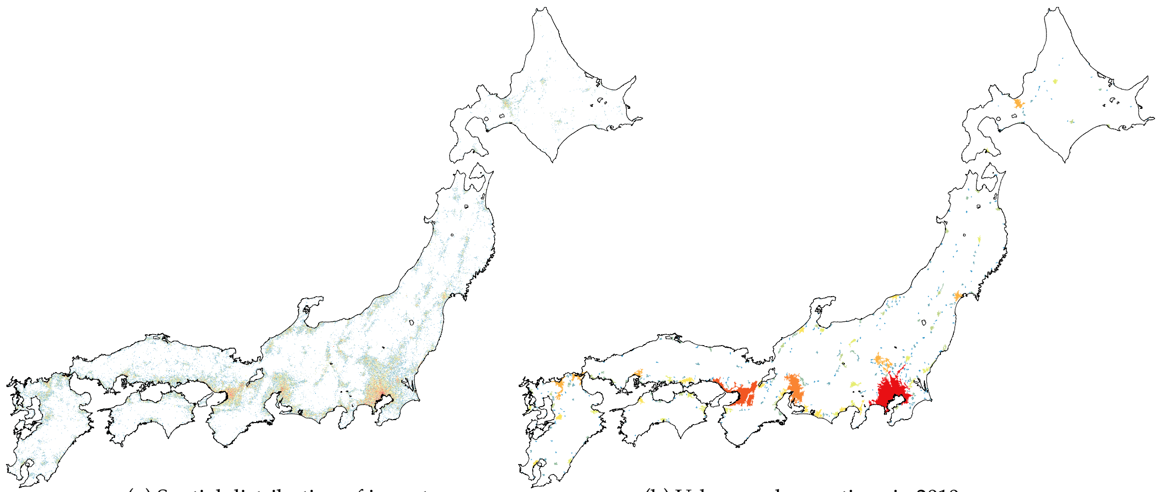
A.3 Locational factors

In this section, the description of UAs and precise definitions for the measures of the local factors discussed in Section 5.3 are given.

UAs – Panels (a) and (b) in Figure A.1 show the spatial distribution of inventors in I and 453 UAs as of 2010, respectively, where the warmer colors in each panel indicate higher population density. Each inventor is assigned to the closest UA if there is any UA within 10 km of their location.

Inventor population – The local population, a_{it}^{INV}, of inventors within a given distance, $\bar{d}$, of the location of inventor i is defined as

$$a_{it}^{\mathrm{INV}} = \left|\left\{ j \in I_t \backslash N_{it} \ : \ d(i,j) < \bar{d} \right\}\right|, \tag{A.1}$$



(a) Spatial distribution of inventors (b) Urban agglomerations in 2010

Figure A.1: Spatial distribution of researchers and UAs

where $d(i, j)$ represents the great-circle distance between inventors i and j (rows 1-4, Table A.3). To evaluate the pure spillover effects, this population excludes the collaborators, N_{it}, of i.[62]

R&D expenditure – Focusing on manufacturing, we first aggregate firm-level R&D expenditure at the industry level according to the three-digit Japanese SIC in each period t. Denote the industry-level R&D expenditure (in million yen) by v_m for each industry $m \in M_t$, where M_t is the set of three-digit manufacturing industries in period t.[63]

Next, from the micro data of the Establishment and Enterprise Census as well as the Economic Census (MIAC, 1996, 2001, 2006; 2009), we find the set of establishments, E_{mt}, in each industry $m \in M_t$ in period t, and compute the employment share, e_{kt}, of each establishment $k \in E_{mt}$ within industry m.

Assuming that the R&D expenditure of each establishment in each industry is proportional to the employment size of the establishment, the value of R&D expenditure of each establishment in period t is approximated by $v_{mt}e_{mt}$. Assuming that the R&D expenditure in the previous period $t-1$ affects the productivity of inventors in the current period t, the R&D around inventor i in period t is given as follows (rows 5-8,

[62]The effects of externalities from the nearby inventors and firms that have been recognized in the literature (e.g., Jaffe et al., 1993; Thompson and Fox-Kean, 2005; Murata et al., 2014; Kerr and Kominers, 2015).

[63]Data on R&D expenditure at the firm level are available for firms with at least four employees for every year from 1997 to 2009 from the Survey of Research and Development. Since we do not have data in 1995 and 1996, the total expenditure in 1997–1999 has been inflated by 1.67 times to obtain the value of R&D expenditure in period 0.

Table A.3):[64]

$$a_{it}^{\mathrm{R\&D}} = \sum_{m \in M_t} \sum_{k \in \{j \in E_m : d(i,j) < \bar{d}\}} v_{m,t-1} e_{k,t-1}. \tag{A.2}$$

$a_{it}^{\mathrm{R\&D}}$ is naturally expected to influence patent development.

Table A.3: Descriptive statistics of the locational factors

Period		(1) 1	(2) 2
(1) inventor population	1km	5,750 (7,225)	5,629 (7,282)
(2)	5km	31,026 (42,143)	30,158 (42,269)
(3)	10km	70,720 (79,277)	66,011 (77,330)
(4)	20km	140,204 (129,401)	127,470 (120,751)
(5) R&D investment	1km	10,454 (78,020)	18,480 (180,284)
(6)	5km	150,581 (338,668)	278,911 (703,381)
(7)	10km	300,256 (466,130)	520,066 (920,505)
(8)	20km	550,420 (584,891)	899,652 (1,098,091)
(9) Manufacturing employment	1km	2,240 (1,505)	6,676 (7,106)
(10)	5km	52,974 (32,395)	76,491 (74,655)
(11)	10km	182,597 (106,414)	212,371 (166,473)
(12)	20km	551,875 (318,789)	509,703 (322,326)
(13) Manufacturing output (in thousand)	1km	21,801,942 (58,182,730)	20,774,589 (83,883,736)
(14)	5km	158,183,183 (129,167,825)	104,957,604 (129,388,708)
(15)	10km	445,908,195 (255,976,915)	317,846,559 (226,259,080)
(16)	20km	1,213,122,353 (626,842,420)	956,808,207 (532,719,932)
(17) Residential population	5km	595,461 (386,442)	615,722 (399,930)
(18)	10km	2,100,541 (1,388,078)	2,156,271 (1,432,171)
(19)	20km	6,386,959 (4,252,098)	6,573,357 (4,416,168)

Numbers are the average values with standard deviations in parentheses.

Manufacturing concentration – Assuming that the employment size and output of an establishment correlate with demand for new knowledge, we proxy the local market size for an invented technology around inventor *i* by the local manufacturing employ-

[64] The R&D expenditure values are obtained from the Survey of Research and Development (1997-2010b) by MIAC and from METI Basic Survey of Japanese Business Structure and Activities (1995-2010) by METI.

ment and output around i:[65]

$$a_{it}^{\mathrm{MNF}_j} = \sum_{k \in \{j \in E_t : d(i,j) < \bar{d}\}} e_{kt} \tag{A.3}$$

where $E_t = \cup_{m \in M_t} E_{mt}$, and e_{kt} represents the total output value (employment) of establishment k for $j = o$ ($j = e$) (rows 9-16, Table A.3).[66]

Residential population – The local residential population is defined as

$$a_{it}^{\mathrm{POP}} = \sum_{k \in \{j \in R : d(i,j) < \bar{d}\}} r_{kt} \tag{A.4}$$

where R represents the set of 1km-by-1km cells covering the relevant location space in Japan; the centroid of each cell is considered to be the representative location of the cell in measuring the distance from the cell; r_{kt} is the residential population in cell $k \in R$ at the beginning of period t (rows 17-19, Table A.3).[67]

B Similarity and difference with linear-in-means models

This section discusses the similarity and difference of instruments between the linear-in-means models of social interactions as in Bramoullé et al. (2009) and our model.

The most relevant similarity is the reflection problem intrinsic to the agent network in both cases, while the most fundamental difference is whether the relevance of the IVs is intrinsic or extrinsic to the network of agents in the model.

In the case of the peer effects in the linear-in-means models, the relevance accrues from the simultaneous equation structure of the model, and thus it is intrinsic to the network. As a consequence, adding degrees of separation in the network is double-edged: the IVs constructed from more distant indirect collaborators can gain exogeneity only at the cost of loosing the relevance. For this reason, the IVs in Bramoullé et al. (2009) are constructed from the exogenous variables of relatively close indirect collaborators in order to retain sufficiently strong relevance. A great advantage in their model is that their IVs formally satisfy the exclusion restriction, provided that the network is exogenous.

In our case, the relevance of the IVs is extrinsic to the inventor network, since it

[65] Another interpretation of a_{it}^{MNF} is the spillover from the manufacturing concentration around i in period t.

[66] The manufacturing employment values are obtained from the Establishment and Enterprise Census for (1996, 2001, 2006) and Economic Census for Business Frame (2009) by MIAC; the manufacturing output values are obtained from the micro data of the Census of Manufacturers (1995, 2000, 2005) and Economic Census for Business Frame (2009) by MIAC.

[67] The residential population in the 1 km-by-1 km cells is available from the Population Census (1995, 2000, 2005) by MIAC.

comes from the similarity in inventor productivity as a result of assortative matching between firms and workers that happened prior to the formation of the inventor network. As a consequence, the relevance is maintained even when the information of the distant indirect collaborators is solely used, as long as the assortative matching affects the indirect collaborators and the targeted inventors simultaneously. That is, the increasing the separation in the network is not double edged. While the endogeneity of the IVs is only virtually (but never completely) eliminated by using sufficiently distant indirect collaborators to construct IVs in our case unlike the case of the linear-in-means models, we instead can allow for the endogenous network formation.

C First-stage regressions for models (4.1) and (4.6)

This section presents the results of the first-stage regressions for the 2SLS IV regressions corresponding to columns 2-5 and 7-10 in Table 7.1 and those in Table 7.3 in tables C.1 and C.2, respectively.

Table C.1: Regression results (Dependent variable: $\ln k_{it}^{D}$)

	Citations				Novelty			
Variables	(1) IV3-5	(2) IV3	(3) IV4	(4) IV5	(5) IV3-5	(6) IV3	(7) IV4	(8) IV5
(1) k_{it}^{D,IV_3}	0.436***	0.453***			0.340***	0.402***		
	(0.0266)	(0.0169)			(0.0146)	(0.0132)		
(2) k_{it}^{D,IV_4}	0.0235*		0.349***		0.0880***		0.347***	
	(0.0128)		(0.0136)		(0.0149)		(0.0149)	
(3) k_{it}^{D,IV_5}	0.00544			0.249***	0.0411*			0.266***
	(0.0409)			(0.0271)	(0.0225)			(0.0250)
(4) $\ln k_{it}$	0.124***	0.124***	0.131***	0.134***	0.156***	0.156***	0.166***	0.176***
	(0.0212)	(0.0210)	(0.0266)	(0.0304)	(0.0126)	(0.0128)	(0.0130)	(0.0145)
(5) $(\ln k_{it})^2$	-0.0491***	-0.0491***	-0.0524***	-0.0545***	-0.0900***	-0.0892***	-0.0934***	-0.0949***
	(0.00741)	(0.00738)	(0.0107)	(0.0126)	(0.00766)	(0.00777)	(0.00967)	(0.0114)
(6) $\ln a_{it}^{\mathrm{INV}}$	0.359***	0.360***	0.390***	0.402***	0.504***	0.515***	0.538***	0.561***
	(0.0787)	(0.0772)	(0.0845)	(0.0859)	(0.0969)	(0.0988)	(0.108)	(0.115)
(7) $\ln a_{it}^{\mathrm{R\&D}}$	0.00240	0.00259	0.00553	0.00979	0.0140	0.0158	0.0178	0.0252
	(0.00909)	(0.00921)	(0.0104)	(0.0113)	(0.0137)	(0.0142)	(0.0166)	(0.0181)
(8) $\ln a_{it}^{\mathrm{MNF}e}$	-0.0668***	-0.0663***	-0.0745***	-0.0759***	-0.0954***	-0.0943***	-0.111***	-0.118***
	(0.0196)	(0.0189)	(0.0244)	(0.0266)	(0.0221)	(0.0197)	(0.0238)	(0.0239)
(9) $\ln a_{it}^{\mathrm{MNF}o}$	0.0227	0.0227	0.0242	0.0242	0.0160	0.0151	0.0149	0.00968
	(0.0207)	(0.0204)	(0.0247)	(0.0246)	(0.0294)	(0.0273)	(0.0320)	(0.0317)
(10) $\ln a_{it}^{\mathrm{POP}}$	1.139	1.143	1.391	1.510	3.041***	3.084***	3.466***	3.699***
	(0.935)	(0.918)	(1.112)	(1.148)	(1.042)	(0.985)	(1.163)	(1.240)
(11) τ_1	0.285***	0.288***	0.333***	0.369***	0.474***	0.504***	0.548***	0.611***
	(0.0211)	(0.0190)	(0.0233)	(0.0288)	(0.0346)	(0.0345)	(0.0382)	(0.0443)
(12) R^2	0.205	0.205	0.183	0.171	0.203	0.201	0.188	0.179
(13) F	443.2	718.4	652.5	84.21	398.6	925.4	541.2	113.2
(14) p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
(15) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) inventor, IPC class and period fixed effects are controlled. (*iii*) ***p<0.01, **p<0.05, * p<0.1

Table C.2: Regression results (Dependent variable: $\ln \Delta n_{it}$)

	Citations				Novelty			
Variables	(1) IV3-5	(2) IV3	(3) IV4	(4) IV5	(5) IV3-5	(6) IV3	(7) IV4	(8) IV5
(1) $\ln \Delta n_{it}^{\mathrm{IV}_3}$	0.244***	0.278***			0.244***	0.278***		
	(0.0212)	(0.0138)			(0.0212)	(0.0138)		
(2) $\ln \Delta n_{it}^{\mathrm{IV}_4}$	0.00997		0.231***		0.00997		0.231***	
	(0.0321)		(0.0304)		(0.0321)		(0.0304)	
(3) $\ln \Delta n_{it}^{\mathrm{IV}_5}$	0.106***			0.208***	0.106***			0.208***
	(0.0339)			(0.0392)	(0.0339)			(0.0392)
(4) $\ln k_{it}$	0.114***	0.116***	0.117***	0.117***	0.114***	0.116***	0.117***	0.117***
	(0.0268)	(0.0266)	(0.0261)	(0.0255)	(0.0268)	(0.0266)	(0.0261)	(0.0255)
(5) $(\ln k_{it})^2$	-0.202***	-0.202***	-0.203***	-0.204***	-0.202***	-0.202***	-0.203***	-0.204***
	(0.00986)	(0.00962)	(0.00971)	(0.00984)	(0.00986)	(0.00962)	(0.00971)	(0.00984)
(6) $\ln a_{it}^{\mathrm{INV}}$	0.232***	0.237***	0.261***	0.273***	0.232***	0.237***	0.261***	0.273***
	(0.0529)	(0.0542)	(0.0600)	(0.0651)	(0.0529)	(0.0542)	(0.0600)	(0.0651)
(7) $\ln a_{it}^{\mathrm{R\&D}}$	0.00704	0.00845	0.00680	0.00712	0.00704	0.00845	0.00680	0.00712
	(0.00751)	(0.00733)	(0.00840)	(0.00874)	(0.00751)	(0.00733)	(0.00840)	(0.00874)
(8) $\ln a_{it}^{\mathrm{MNF}_e}$	-0.0329	-0.0326*	-0.0412*	-0.0427*	-0.0329	-0.0326*	-0.0412*	-0.0427*
	(0.0200)	(0.0194)	(0.0220)	(0.0240)	(0.0200)	(0.0194)	(0.0220)	(0.0240)
(9) $\ln a_{it}^{\mathrm{MNF}_o}$	0.00479	0.00514	0.00793	0.00868	0.00479	0.00514	0.00793	0.00868
	(0.0142)	(0.0142)	(0.0155)	(0.0170)	(0.0142)	(0.0142)	(0.0155)	(0.0170)
(10) $\ln a_{it}^{\mathrm{POP}}$	1.053**	1.106**	1.210**	1.338**	1.053**	1.106**	1.210**	1.338**
	(0.506)	(0.539)	(0.585)	(0.585)	(0.506)	(0.539)	(0.585)	(0.585)
(11) τ_1	-0.167***	-0.146***	-0.139***	-0.131***	-0.167***	-0.146***	-0.139***	-0.131***
	(0.0206)	(0.0187)	(0.0252)	(0.0263)	(0.0206)	(0.0187)	(0.0252)	(0.0263)
(12) R^2	0.197	0.196	0.190	0.189	0.197	0.196	0.190	0.189
(13) F	142.2	406.1	57.40	28.24	142.2	406.1	57.40	28.24
(14) p-value	0.000	0.000	0.000	3.44e-07	0.000	0.000	0.000	3.44e-07
(15) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) inventor, IPC class and period fixed effects are controlled. (*iii*) ***p<0.01, **p<0.05, * p<0.1

D Results for model (4.5) with $m = p$

Table D.1 shows the second stage regression results for model (4.5) with $m = p$.

Table D.1: Regression results (Dependent variable: $\ln y_{it}^{p}$)

	Citations					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln k_{it}^{D}$	0.135***	0.260***	0.259***	0.272***	0.270***	0.0451***	0.114***	0.110***	0.132***	0.130***
	(0.0103)	(0.0283)	(0.0290)	(0.0271)	(0.0373)	(0.00521)	(0.0214)	(0.0246)	(0.0170)	(0.0416)
(2) $\ln k_{it}$	0.0999***	0.0825***	0.0826***	0.0808***	0.0811***	0.111***	0.0982***	0.0989***	0.0949***	0.0952***
	(0.0110)	(0.00729)	(0.00725)	(0.00844)	(0.0125)	(0.0144)	(0.0112)	(0.0106)	(0.0125)	(0.0194)
(3) $(\ln k_{it})^2$	-0.0835***	-0.0765***	-0.0765***	-0.0758***	-0.0759***	-0.0868***	-0.0803***	-0.0807***	-0.0787***	-0.0788***
	(0.00921)	(0.00761)	(0.00759)	(0.00796)	(0.00884)	(0.0104)	(0.00821)	(0.00800)	(0.00869)	(0.0115)
(4) $\ln a_{it}^{\mathrm{INV}}$	0.207***	0.153**	0.153**	0.148**	0.149***	0.238***	0.196***	0.199***	0.185***	0.186***
	(0.0678)	(0.0752)	(0.0755)	(0.0700)	(0.0574)	(0.0651)	(0.0659)	(0.0683)	(0.0612)	(0.0389)
(5) $\ln a_{it}^{\mathrm{R\&D}}$	0.0298***	0.0281***	0.0281***	0.0279***	0.0280***	0.0302***	0.0280***	0.0282***	0.0275***	0.0276***
	(0.0105)	(0.00925)	(0.00927)	(0.00892)	(0.00867)	(0.0112)	(0.0101)	(0.0102)	(0.00967)	(0.00885)
(6) $\ln a_{it}^{\mathrm{MNF}_e}$	-0.0122*	-0.00309	-0.00313	-0.00222	-0.00236	-0.0166**	-0.00823	-0.00872	-0.00604	-0.00626
	(0.00702)	(0.00616)	(0.00611)	(0.00725)	(0.00869)	(0.00772)	(0.00745)	(0.00756)	(0.00718)	(0.00950)
(7) $\ln a_{it}^{\mathrm{MNF}_o}$	-0.000286	-0.00341	-0.00340	-0.00371	-0.00366	0.00274	0.00217	0.00220	0.00202	0.00203
	(0.00354)	(0.00438)	(0.00438)	(0.00435)	(0.00426)	(0.00420)	(0.00344)	(0.00345)	(0.00339)	(0.00343)
(8) $\ln a_{it}^{\mathrm{POP}}$	0.133	-0.0794	-0.0785	-0.0997	-0.0965	0.183	-0.0932	-0.0768	-0.165	-0.158
	(0.461)	(0.420)	(0.420)	(0.409)	(0.387)	(0.535)	(0.509)	(0.517)	(0.495)	(0.458)
(9) τ_1	0.126***	0.0708***	0.0710***	0.0655***	0.0663***	0.153***	0.103***	0.106***	0.0898***	0.0911***
	(0.0183)	(0.0215)	(0.0217)	(0.0199)	(0.0223)	(0.0179)	(0.0236)	(0.0254)	(0.0206)	(0.0342)
(10) R^2	0.102	0.087	0.087	0.084	0.084	0.089	0.078	0.079	0.072	0.072
(11) Hansen J p-val.		0.878					0.177			
(12) 1st stage F		727.1	2178	1080	509.6		557.6	1590	918.7	471.4
(13) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) ***p<0.01, ** p<0.05, * p<0.1.

E Other robustness analyses

E.1 Alternative productivity measures

This section presents the regression results for (4.1) and (4.6) under the four alternative measures of inventor productivity, where the output, g_j, of patent j in (2.1) is given by (*i*) the cited count within five years from publication, (*ii*) technological novelty based on the IPC subclass, (*iii*) count of patent claims; or (*iv*) count of patents, i.e., $g_j = 1$ for all j. Table E.1 shows the descriptive statistics for productivities and differentiated knowledge of collaborators under these measures.

Table E.1: Descriptive statistics of knowledge and productivity variables

Unit of productivity		**Citations (5 years)**		**Novelty (IPC subclass)**		**Claim counts**		**Patent counts**	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Period		1	2	1	2	1	2	1	2
(1) Output of a patent	g_{jt}	1.789 (3.676)	1.595 (4.186)	0.000 (0.002)	0.000 (0.003)	7.231 (9.555)	8.906 (81.53)	1.000 (0.000)	1.000 (0.000)
(2) Productivity of an inventor	y_{it}	9.369 (22.02)	5.597 (163.73)	0.001 (0.004)	0.000 (0.003)	36.67 (109.14)	40.89 (4173.48)	4.824 (7.749)	3.099 (5.936)
(3) Pairwise productivity of an inventor	y_{it}	1.622 (3.730)	1.838 (175.40)	0.000 (0.001)	0.000 (0.001)	6.682 (88.27)	25.48 (4478.60)	0.894 (1.911)	0.677 (1.739)
(4) Avg. diff. knowledge of collaborators	k_{it}^D	1.666 (8.940)	1.186 (5.391)	0.008 (0.043)	0.005 (0.034)	6.579 (49.51)	5.647 (25.86)	0.874 (4.699)	0.748 (3.057)

Numbers are the average values with standard deviations in parentheses.

Tables E.2 and E.3 present the results from the second-stage regressions for models (4.1) and (4.6), respectively. Under each alternative measure, the tables show the OLS and the IV results, where the IVs for $\ln k_{it}^D$ and $\ln \Delta n_{it}$ are constructed by using all indirect collaborators for $\ell = 3, 4$ and 5, since the result is similar if only one of them is used (just like in our baseline results).

Table E.2: Regression results for (4.1) under alternative productivity measures

	Citations (5 years)		Novelty (IPC subclass)		Claim count		Patent count	
Variables	(1) OLS	(2) IV3-5	(3)OLS	(4) IV3-5	(5) OLS	(6) IV3-5	(7)OLS	(8) IV3-5
(1) $\ln k_{it}^D$	0.163***	0.282***	0.179***	0.323***	0.198***	0.339***	0.163***	0.326***
	(0.0109)	(0.0276)	(0.00882)	(0.0347)	(0.0133)	(0.0383)	(0.0112)	(0.0304)
(2) $\ln k_{it}$	0.116***	0.0991***	0.0586***	0.0394**	0.140***	0.115***	0.0981***	0.0775***
	(0.0155)	(0.0117)	(0.0199)	(0.0186)	(0.0135)	(0.0152)	(0.0114)	(0.00723)
(3) $(\ln k_{it})^2$	-0.0887***	-0.0820***	-0.159***	-0.146***	-0.0953***	-0.0858***	-0.0826***	-0.0742***
	(0.01000)	(0.00892)	(0.0184)	(0.0121)	(0.00337)	(0.00338)	(0.00911)	(0.00715)
(4) $\ln a_{it}^{\mathrm{INV}}$	0.167***	0.118*	0.232***	0.153**	0.212***	0.142**	0.187***	0.109
	(0.0539)	(0.0606)	(0.0644)	(0.0707)	(0.0559)	(0.0579)	(0.0658)	(0.0734)
(5) $\ln a_{it}^{\mathrm{R\&D}}$	0.0269***	0.0254***	0.0415***	0.0379***	0.0276***	0.0251***	0.0290***	0.0265***
	(0.00744)	(0.00650)	(0.0127)	(0.0106)	(0.0101)	(0.00865)	(0.00971)	(0.00767)
(6) $\ln a_{it}^{\mathrm{MNF}_e}$	0.0188***	0.0269***	0.00689	0.0214**	0.0148**	0.0274***	-0.00502	0.0120**
	(0.00566)	(0.00465)	(0.0107)	(0.00993)	(0.00658)	(0.00542)	(0.00639)	(0.00504)
(7) $\ln a_{it}^{\mathrm{MNF}_o}$	0.00857	0.00546	-0.00516	-0.00643	0.0127**	0.00860	0.000798	-0.00151
	(0.00616)	(0.00835)	(0.00456)	(0.00637)	(0.00508)	(0.00604)	(0.00344)	(0.00586)
(8) $\ln a_{it}^{\mathrm{POP}}$	-0.435	-0.626	0.394	0.00176	0.594	0.154	0.0162	-0.331
	(0.527)	(0.494)	(0.533)	(0.466)	(0.478)	(0.484)	(0.450)	(0.409)
(9) τ_1	0.272***	0.214***	0.433***	0.327***	0.122***	0.0788***	0.133***	0.0801***
	(0.0175)	(0.0156)	(0.0247)	(0.0352)	(0.0236)	(0.0230)	(0.0160)	(0.0155)
(10) R^2	0.165		0.236		0.105		0.107	
(11) Hansen J p-val.		0.845		0.184		0.399		0.629
(12) 1st stage F		712.6		500.8		889.2		774.7
(13) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Table E.3: Regression results for (4.6) under alternative productivity measures

	Citations (5 years)		Novelty (IPC subclass)		Claim count		Patent count	
Variables	(1) OLS	(2) IV3-5	(3)OLS	(4) IV3-5	(5) OLS	(6) IV3-5	(7)OLS	(8) IV3-5
(1) $\ln \Delta n_{it}$	0.107***	1.383***	0.174***	1.427***	0.110***	1.525***	0.0851***	1.318***
	(0.00587)	(0.0683)	(0.00853)	(0.0920)	(0.00630)	(0.0540)	(0.00595)	(0.0512)
(2) $\ln k_{it}$	0.137***	-0.0166	0.122***	0.000294	0.166***	-0.00414	0.116***	-0.0324
	(0.0452)	(0.0688)	(0.0467)	(0.0914)	(0.0234)	(0.0526)	(0.0307)	(0.0563)
(3) $(\ln k_{it})^2$	-0.0366**	0.225***	-0.0551	0.242***	-0.0456***	0.244***	-0.0343***	0.218***
	(0.0160)	(0.0191)	(0.0340)	(0.0389)	(0.0103)	(0.0200)	(0.0131)	(0.0200)
(4) $\ln a_{it}^{\mathrm{INV}}$	0.365***	-0.00994	0.479***	0.144**	0.445***	0.0288	0.450***	0.0876
	(0.0965)	(0.0387)	(0.0967)	(0.0611)	(0.131)	(0.0561)	(0.0913)	(0.0534)
(5) $\ln a_{it}^{\mathrm{R\&D}}$	0.0129	0.000127	0.0239	0.00825	0.0186	0.00446	0.0156	0.00330
	(0.0106)	(0.00535)	(0.0151)	(0.00960)	(0.0129)	(0.00568)	(0.0140)	(0.00545)
(6) $\ln a_{it}^{\mathrm{MNF}_e}$	-0.0666***	-0.00955	-0.0990***	-0.0673***	-0.0916***	-0.0283	-0.106***	-0.0505***
	(0.0212)	(0.0150)	(0.0220)	(0.0176)	(0.0233)	(0.0203)	(0.0219)	(0.0139)
(7) $\ln a_{it}^{\mathrm{MNF}_o}$	0.0221	0.00871	0.00282	-0.0220*	0.0225	0.00768	0.00888	-0.00401
	(0.0204)	(0.0112)	(0.0265)	(0.0130)	(0.0257)	(0.0132)	(0.0253)	(0.00891)
(8) $\ln a_{it}^{\mathrm{POP}}$	1.238	-0.698	2.201**	-0.0572	2.924***	0.778	1.808*	-0.0609
	(1.058)	(1.278)	(0.879)	(1.056)	(0.886)	(1.167)	(0.951)	(1.000)
(9) τ_1	0.459***	0.560***	0.674***	0.522***	0.287***	0.398***	0.301***	0.398***
	(0.0289)	(0.0539)	(0.0265)	(0.0398)	(0.0262)	(0.0322)	(0.0215)	(0.0319)
(10) R^2	0.177		0.217		0.089		0.111	
(11) Hansen J p-val.		0.254		0.297		0.245		0.251
(12) 1st stage F		237.7		251.2		237.7		237.7
(13) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

E.2 Alternative definition of collaborators' knowledge

Table E.4 shows the regression results for (4.1) under citation- and novelty-adjusted productivity measures in columns 1–2 and 3–4, respectively, and those for (4.6) in columns 5–6 when k_{it}^{D} is defined in terms of IPC subgroups as described in Section 8.3. For each specification, we compare the OLS and IV results, where the latter are shown only for the case in which IVs are constructed by using all the third-, forth- and fifth-indirect collaborators, since the result is similar, even if only either of the third-, forth- or fifth-indirect collaborators were used.

In all the specifications, the first-stage F values are reasonably large, so that the relevance appears to be strong as in the baseline case. In terms of the Hansen (1982)'s J-test, there is no evidence against the exogeneity of the instruments.

Table E.4: Regression results with knowledge in terms of IPC subgroups

	(4.1) Dependent variable : $\ln y_{it}$				(4.6) Dependent variable : $\ln \Delta n_{it}$	
	Citations		Novelty			
Variables	(1) OLS	(2) IV3-5	(3)OLS	(4) IV3-5	(5) OLS	(6) IV3-5
(1) $\ln k_{it}^{D}$	0.169***	0.604***	0.287***	1.331***		
	(0.0431)	(0.118)	(0.0597)	(0.244)		
(2) $\ln \Delta n_{it}$					-0.0210***	0.368***
					(0.00386)	(0.0160)
(3) $\ln k_{it}$	0.128***	0.111***	0.167***	0.126***	0.0432***	-0.00226
	(0.0201)	(0.0166)	(0.0163)	(0.0246)	(0.00642)	(0.0157)
(4) $(\ln k_{it})^2$	-0.0983***	-0.0936***	-0.208***	-0.197***	-0.0179***	0.0611***
	(0.0116)	(0.0107)	(0.0106)	(0.00802)	(0.00184)	(0.00651)
(5) $\ln a_{it}^{\mathrm{INV}}$	0.230***	0.189***	0.386***	0.287***	0.105***	-0.00729
	(0.0596)	(0.0610)	(0.103)	(0.109)	(0.0277)	(0.0213)
(6) $\ln a_{it}^{\mathrm{R\&D}}$	0.0281***	0.0270***	0.0448***	0.0422***	0.00259	-0.00116
	(0.00879)	(0.00771)	(0.0170)	(0.0143)	(0.00380)	(0.00188)
(7) $\ln a_{it}^{\mathrm{MNF}_e}$	0.00313	0.00820	-0.0293**	-0.0171	-0.0148***	0.00371
	(0.00823)	(0.00705)	(0.0114)	(0.0110)	(0.00461)	(0.00650)
(8) $\ln a_{it}^{\mathrm{MNF}_o}$	0.0113***	0.0112**	-0.00369	-0.00392	-0.00126	-0.00551
	(0.00431)	(0.00555)	(0.00661)	(0.00810)	(0.00579)	(0.00378)
(9) $\ln a_{it}^{\mathrm{POP}}$	-0.375	-0.727	1.238**	0.392	0.903***	0.299
	(0.627)	(0.635)	(0.535)	(0.587)	(0.200)	(0.247)
(10) τ_1	0.279***	0.256***	0.399***	0.344***	0.0453***	0.0754***
	(0.0195)	(0.0176)	(0.0339)	(0.0387)	(0.00639)	(0.00867)
(11) R^2	0.132		0.169		0.018	
(12) Hansen J p-val.		0.931		0.109		0.235
(13) 1st stage F		328.8		328.8		235.2
(14) #Obs.	113,454	113,454	113,454	113,454	92,098	92,098

(*i*) standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

E.3 Results for the alternative radiuses for locational factors

This section presents the results from the second-stage regressions for (4.1) in Section 7.1 and (4.6) in Section 7.3 under the alternative radius values for the local factors defined in Section 5.3 in stables E.5 and E.6 (E.7 and E.8), respectively for quality-adjusted (novelty-adjusted) productivity.

One can see that the choice of radius values for the local factors does not alter the qualitative results obtained in the baseline setup shown in tables 7.1 and 7.3 in Section 7 regarding the effect of collaborators' differentiated knowledge and that of the knowledge stock of an inventor on their productivity as well as the role of the collaborator recombination in the size of collaborators' differentiated knowledge. The values of the estimated coefficients for the endogenous variables, $\ln k_{it}^{D}$ and $\ln\Delta n_{it}$, as well as those for the knowledge stock, $\ln k_{it}$ and $(\ln k_{it})^2$, appear to be stable in all cases.

Table E.5: Regression results (Dependent variable: $\ln y_{it}$)

	Citations (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln k_{it}^{D}$	0.293***	0.296***	0.294***	0.286***	0.288***	0.288***	0.286***
	(0.0225)	(0.0193)	(0.0214)	(0.0254)	(0.0239)	(0.0242)	(0.0254)
(2) $\ln k_{it}$	0.0923***	0.0893***	0.0889***	0.0931***	0.0917***	0.0911***	0.0931***
	(0.0124)	(0.0135)	(0.0141)	(0.0119)	(0.0121)	(0.0121)	(0.0119)
(3) $(\ln k_{it})^2$	-0.0814***	-0.0803***	-0.0802***	-0.0820***	-0.0815***	-0.0816***	-0.0820***
	(0.00886)	(0.00903)	(0.00911)	(0.00868)	(0.00879)	(0.00866)	(0.00868)
(4) $\ln a_{it}^{\mathrm{INV}}$							
1km				0.117*	0.126**	0.125*	0.117*
				(0.0633)	(0.0634)	(0.0709)	(0.0633)
5km	0.162***						
	(0.0624)						
10km		0.0886					
		(0.108)					
20km			0.127				
			(0.135)				
(5) $\ln a_{it}^{\mathrm{R\&D}}$							
1km	0.0267***	0.0260***	0.0270***				0.0256***
	(0.00611)	(0.00734)	(0.00838)				(0.00679)
5km				0.0256***			
				(0.00679)			
10km					0.0294***		
					(0.0106)		
20km						0.0314***	
						(0.00865)	
(6) $\ln a_{it}^{\mathrm{MNF_e}}$							
1km	0.0113	0.0176**	0.0209***	0.0240***	0.0277***	0.0202***	
	(0.00825)	(0.00799)	(0.00487)	(0.00438)	(0.00597)	(0.00427)	
5km							0.0240***
							(0.00438)
(7) $\ln a_{it}^{\mathrm{MNF_o}}$							
1km	0.00492	0.00563	0.00650	0.00522	0.00534	0.00588	0.00522
	(0.00863)	(0.00821)	(0.00880)	(0.00804)	(0.00667)	(0.00722)	(0.00804)
(8) $\ln a_{it}^{\mathrm{POP}}$							
1km	-0.624	-0.628	-0.628	-0.660	-0.939*	-0.522	-0.660
	(0.472)	(0.517)	(0.546)	(0.490)	(0.497)	(0.504)	(0.490)
(9) τ_1	0.165***	0.166***	0.163***	0.173***	0.164***	0.166***	0.173***
	(0.0143)	(0.0174)	(0.0190)	(0.0150)	(0.0196)	(0.0117)	(0.0150)
(10) H. J p-value	0.952	0.972	0.974	0.928	0.938	0.878	0.928
(11) F	768.5	775.2	758.3	727.1	734	733.4	727.1
(12) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) ***p<0.01, ** p<0.05, * p<0.1.

Table E.5: Regression results continued (Dependent variable: $\ln y_{it}$)

	Citations (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln k_{it}^{D}$	0.284***	0.286***	0.284***	0.287***	0.280***	0.285***	0.280***
	(0.0249)	(0.0248)	(0.0274)	(0.0248)	(0.0280)	(0.0249)	(0.0264)
(2) $\ln k_{it}$	0.0932***	0.0923***	0.0913***	0.0931***	0.0907***	0.0921***	0.0903***
	(0.0115)	(0.0105)	(0.0121)	(0.0119)	(0.0117)	(0.0107)	(0.0117)
(3) $(\ln k_{it})^2$	-0.0822***	-0.0820***	-0.0817***	-0.0821***	-0.0817***	-0.0815***	-0.0809***
	(0.00864)	(0.00848)	(0.00867)	(0.00857)	(0.00861)	(0.00834)	(0.00861)
(4) $\ln a_{it}^{\mathrm{INV}}$							
1km	0.117*	0.124**	0.119*	0.118*	0.115*	0.116*	0.108*
	(0.0634)	(0.0599)	(0.0609)	(0.0608)	(0.0632)	(0.0654)	(0.0598)
(5) $\ln a_{it}^{\mathrm{R\&D}}$							
1km	0.0251***	0.0240***	0.0209***	0.0274***	0.0177***	0.0275***	0.0271***
	(0.00780)	(0.00911)	(0.00498)	(0.00570)	(0.00635)	(0.00634)	(0.00530)
(6) $\ln a_{it}^{\mathrm{MNF_e}}$							
1km			0.0488***	0.0155	0.0283***	0.0218***	0.0217***
			(0.0159)	(0.0157)	(0.00899)	(0.00515)	(0.00695)
10km	-0.0245						
	(0.0165)						
20km		-0.105					
		(0.0783)					
(7) $\ln a_{it}^{\mathrm{MNF_o}}$							
1km	-0.00201	-0.00280				0.00736	0.0102*
	(0.0119)	(0.0129)				(0.00539)	(0.00619)
5km			0.0573***				
			(0.0199)				
10km				-0.0122			
				(0.0481)			
20km					0.137***		
					(0.0483)		
(8) $\ln a_{it}^{\mathrm{POP}}$							
5km						0.0295	
						(0.276)	
10km							0.666
							(0.547)
20km	-0.592	-0.491	-0.217	-0.806	0.232		
	(0.562)	(0.542)	(0.533)	(0.749)	(0.592)		
(9) τ_1	0.177***	0.190***	0.160***	0.178***	0.162***	0.189***	0.203***
	(0.0153)	(0.0173)	(0.0132)	(0.0159)	(0.0162)	(0.0213)	(0.0271)
(10) H. J p-value	0.943	0.944	0.875	0.934	0.846	0.920	0.935
(11) F	721.8	728.7	728.5	728.6	722.6	729.1	709.6
(12) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Table E.6: Regression results (Dependent variable: $\ln k_{it}^{D}$)

	Citations (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln \Delta n_{it}$	1.381*** (0.0616)	1.403*** (0.0560)	1.404*** (0.0571)	1.372*** (0.0629)	1.378*** (0.0583)	1.372*** (0.0634)	1.372*** (0.0629)
(2) $\ln k_{it}$	-0.0243 (0.0660)	-0.0279 (0.0671)	-0.0257 (0.0671)	-0.0220 (0.0669)	-0.0212 (0.0673)	-0.0225 (0.0675)	-0.0220 (0.0669)
(3) $(\ln k_{it})^2$	0.226*** (0.0184)	0.231*** (0.0200)	0.230*** (0.0216)	0.223*** (0.0197)	0.224*** (0.0207)	0.223*** (0.0199)	0.223*** (0.0197)
(4) $\ln a_{it}^{\mathrm{INV}}$							
1km				0.0138 (0.0426)	0.0109 (0.0473)	0.0144 (0.0418)	0.0138 (0.0426)
5km	-0.0460 (0.0699)						
10km		-0.287*** (0.100)					
20km			-0.338** (0.171)				
(5) $\ln a_{it}^{\mathrm{R\&D}}$							
1km	0.000665 (0.00443)	0.00203 (0.00617)	-0.000853 (0.00529)				0.000705 (0.00478)
5km				0.000705 (0.00478)			
10km					-0.0150* (0.00821)		
20km						0.00480 (0.0109)	
(6) $\ln a_{it}^{\mathrm{MNFe}}$							
1km	-0.0106 (0.0167)	0.00436 (0.0151)	-0.00815 (0.0141)	-0.0139 (0.0147)	-0.0161 (0.0175)	-0.0144 (0.0146)	
5km							-0.0139 (0.0147)
(7) $\ln a_{it}^{\mathrm{MNFo}}$							
1km	0.00833 (0.00995)	0.00795 (0.00980)	0.00566 (0.0111)	0.00814 (0.00992)	0.0117 (0.0112)	0.00741 (0.00973)	0.00814 (0.00992)
(8) $\ln a_{it}^{\mathrm{POP}}$							
20km	-0.539 (1.231)	-0.458 (1.061)	-0.471 (1.023)	-0.552 (1.229)	-0.688 (1.247)	-0.464 (1.179)	-0.552 (1.229)
(9) τ_1	0.520*** (0.0550)	0.555*** (0.0441)	0.557*** (0.0358)	0.514*** (0.0504)	0.494*** (0.0552)	0.520*** (0.0471)	0.514*** (0.0504)
(10) H. J p-value	0.253	0.243	0.251	0.255	0.253	0.254	0.255
(11) F	266	258.3	249.5	237.7	238	239.6	237.7
(12) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) $^{***}p$<0.01, ** p<0.05, * p<0.1.

Table E.6: Regression results continued (Dependent variable: $\ln k_{it}^{D}$)

	Citations (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln \Delta n_{it}$	1.371***	1.377***	1.379***	1.381***	1.375***	1.366***	1.366***
	(0.0601)	(0.0652)	(0.0631)	(0.0595)	(0.0628)	(0.0614)	(0.0537)
(2) $\ln k_{it}$	-0.0228	-0.0232	-0.0212	-0.0205	-0.0228	-0.0233	-0.0227
	(0.0666)	(0.0661)	(0.0664)	(0.0648)	(0.0659)	(0.0652)	(0.0647)
(3) $(\ln k_{it})^2$	0.223***	0.224***	0.224***	0.225***	0.224***	0.223***	0.223***
	(0.0206)	(0.0194)	(0.0194)	(0.0197)	(0.0192)	(0.0185)	(0.0204)
(4) $\ln a_{it}^{\mathrm{INV}}$							
1km	0.0148	0.0176	0.0124	0.0220	0.0127	0.0117	0.0120
	(0.0421)	(0.0429)	(0.0443)	(0.0440)	(0.0434)	(0.0446)	(0.0491)
(5) $\ln a_{it}^{\mathrm{R\&D}}$							
1km	0.000500	9.21e-05	0.00515	0.0153**	0.000744	0.00240	0.00211
	(0.00483)	(0.00500)	(0.00487)	(0.00729)	(0.00518)	(0.00631)	(0.00686)
(6) $\ln a_{it}^{\mathrm{MNF_e}}$							
1km			-0.0372	-0.0761**	-0.0190*	-0.0183	-0.0154
			(0.0268)	(0.0386)	(0.00986)	(0.0151)	(0.0163)
10km	-0.0345						
	(0.0464)						
20km		-0.0559					
		(0.0614)					
(7) $\ln a_{it}^{\mathrm{MNF_o}}$							
1km	0.00762	0.00926				0.0107	0.0103
	(0.00909)	(0.00726)				(0.00763)	(0.00854)
5km			-0.0337				
			(0.0369)				
10km				-0.152*			
				(0.0809)			
20km					0.0171		
					(0.0579)		
(8) $\ln a_{it}^{\mathrm{POP}}$							
5km						0.135	
						(0.450)	
10km							0.111
							(1.355)
20km	-0.566	-0.524	-0.954	-1.760	-0.529		
	(1.298)	(1.304)	(1.278)	(1.340)	(1.368)		
(9) τ_1	0.516***	0.522***	0.529***	0.553***	0.517***	0.529***	0.529***
	(0.0531)	(0.0574)	(0.0503)	(0.0535)	(0.0512)	(0.0384)	(0.0501)
(10) H. J p-value	0.256	0.252	0.258	0.258	0.256	0.258	0.252
(11) F	238.6	241.1	238.8	239.2	238.4	242.9	230.5
(12) #Obs	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Table E.7: Regression results (Dependent variable: $\ln y_{it}$)

	Novelty (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln k^D_{it}$	0.355***	0.358***	0.351***	0.344***	0.344***	0.345***	0.344***
	(0.0290)	(0.0303)	(0.0312)	(0.0310)	(0.0285)	(0.0306)	(0.0310)
(2) $\ln k_{it}$	0.112***	0.108***	0.107***	0.114***	0.113***	0.111***	0.114***
	(0.0227)	(0.0211)	(0.0204)	(0.0228)	(0.0219)	(0.0231)	(0.0228)
(3) $(\ln k_{it})^2$	-0.176***	-0.175***	-0.175***	-0.178***	-0.177***	-0.177***	-0.178***
	(0.00650)	(0.00698)	(0.00731)	(0.00594)	(0.00608)	(0.00621)	(0.00594)
(4) $\ln a^{\mathrm{INV}}_{it}$							
1km				0.200**	0.213**	0.212**	0.200**
				(0.0939)	(0.0957)	(0.104)	(0.0939)
5km	0.256***						
	(0.0789)						
10km		0.223					
		(0.149)					
20km			0.397**				
			(0.155)				
(5) $\ln a^{\mathrm{R\&D}}_{it}$							
1km	0.0381***	0.0366***	0.0395***				0.0364***
	(0.0116)	(0.0124)	(0.0137)				(0.0127)
5km				0.0364***			
				(0.0127)			
10km					0.0399**		
					(0.0189)		
20km						0.0492***	
						(0.0154)	
(6) $\ln a^{\mathrm{MNF_e}}_{it}$							
1km	-0.00656	-0.00148	0.00538	0.0132	0.0181	0.00757	
	(0.00776)	(0.00851)	(0.00928)	(0.00989)	(0.0126)	(0.0110)	
5km							0.0132
							(0.00989)
(7) $\ln a^{\mathrm{MNF_o}}_{it}$							
1km	-0.00532	-0.00412	-0.00151	-0.00512	-0.00445	-0.00496	-0.00512
	(0.00882)	(0.00790)	(0.00931)	(0.00721)	(0.00467)	(0.00493)	(0.00721)
(8) $\ln a^{\mathrm{POP}}_{it}$							
20km	0.112	0.0752	0.0741	0.0701	-0.340	0.361	0.0701
	(0.414)	(0.426)	(0.489)	(0.415)	(0.377)	(0.443)	(0.415)
(9) τ_1	0.158***	0.150***	0.137***	0.173***	0.159***	0.171***	0.173***
	(0.0376)	(0.0377)	(0.0338)	(0.0382)	(0.0366)	(0.0426)	(0.0382)
(10) H. J p-value	0.663	0.619	0.642	0.768	0.823	0.782	0.768
(11) F	588.2	593.8	568.5	557.6	564.3	563.9	557.6
(12) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) ***p<0.01, ** p<0.05, * p<0.1.

Table E.7: Regression results continued (Dependent variable: $\ln y_{it}$)

	Novelty (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln k^D_{it}$	0.338*** (0.0300)	0.344*** (0.0299)	0.342*** (0.0301)	0.345*** (0.0313)	0.340*** (0.0303)	0.343*** (0.0289)	0.332*** (0.0302)
(2) $\ln k_{it}$	0.114*** (0.0225)	0.113*** (0.0235)	0.112*** (0.0233)	0.114*** (0.0226)	0.111*** (0.0235)	0.113*** (0.0237)	0.109*** (0.0229)
(3) $(\ln k_{it})^2$	-0.178*** (0.00592)	-0.178*** (0.00566)	-0.177*** (0.00622)	-0.178*** (0.00596)	-0.177*** (0.00618)	-0.177*** (0.00549)	-0.177*** (0.00558)
(4) $\ln a^{\mathrm{INV}}_{it}$	0.203** (0.0917)	0.212** (0.0878)	0.202** (0.0912)	0.199** (0.0893)	0.195** (0.0950)	0.198** (0.100)	0.180** (0.0886)
(5) $\ln a^{\mathrm{R\&D}}_{it}$	0.0356** (0.0147)	0.0341** (0.0160)	0.0285*** (0.00994)	0.0355*** (0.0107)	0.0252** (0.0120)	0.0367*** (0.0122)	0.0353*** (0.00869)
(6) $\ln a^{\mathrm{MNF_e}}_{it}$							
1km			0.0557*** (0.0202)	0.0182 (0.0382)	0.0276*** (0.0101)	0.00800 (0.0126)	0.0110 (0.00690)
10km	-0.0728** (0.0294)						
20km		-0.169 (0.106)					
(7) $\ln a^{\mathrm{MNF_o}}_{it}$							
1km	-0.0147** (0.00740)	-0.0128 (0.00967)				-0.00360 (0.00616)	0.00286 (0.00574)
5km			0.0767*** (0.0259)				
10km				0.00257 (0.0890)			
20km					0.168*** (0.0404)		
(8) $\ln a^{\mathrm{POP}}_{it}$							
5km						0.248 (0.331)	
10km							1.870*** (0.672)
20km	0.155 (0.500)	0.286 (0.457)	0.803* (0.474)	0.144 (0.885)	1.293** (0.535)		
(9) τ_1	0.184*** (0.0398)	0.199*** (0.0322)	0.150*** (0.0414)	0.169*** (0.0358)	0.153*** (0.0424)	0.177*** (0.0299)	0.214*** (0.0355)
(10) H. J p-value	0.776	0.654	0.850	0.729	0.842	0.809	0.767
(11) F	550.6	563.4	555.8	557.7	552.2	562.4	537.2
(12) #Obs.	116,928	116,928	116,928	116,928	116,928	116,928	116,928

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Table E.8: Regression results (Dependent variable: $\ln k_{it}^{D}$)

	Novelty (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln \Delta n_{it}$	1.743***	1.764***	1.743***	1.722***	1.730***	1.727***	1.722***
	(0.0748)	(0.0706)	(0.0784)	(0.0847)	(0.0799)	(0.0815)	(0.0847)
(2) $\ln k_{it}$	-0.0355	-0.0397	-0.0359	-0.0313	-0.0310	-0.0342	-0.0313
	(0.0635)	(0.0655)	(0.0659)	(0.0638)	(0.0647)	(0.0654)	(0.0638)
(3) $(\ln k_{it})^2$	0.266***	0.271***	0.266***	0.261***	0.262***	0.262***	0.261***
	(0.0226)	(0.0241)	(0.0243)	(0.0233)	(0.0250)	(0.0241)	(0.0233)
(4) $\ln a_{it}^{\mathrm{INV}}$							
1km				0.0800	0.0820	0.0849	0.0800
				(0.103)	(0.112)	(0.106)	(0.103)
5km	0.0170						
	(0.132)						
10km		-0.208					
		(0.150)					
20km			0.00578				
			(0.208)				
(5) $\ln a_{it}^{\mathrm{R\&D}}$							
1km	0.0177**	0.0186*	0.0177**				0.0172*
	(0.00874)	(0.00989)	(0.00861)				(0.00896)
5km				0.0172*			
				(0.00896)			
10km					-0.00246		
					(0.0159)		
20km						0.0285	
						(0.0177)	
(6) $\ln a_{it}^{\mathrm{MNF_e}}$							
1km	-0.0451***	-0.0307*	-0.0441**	-0.0436**	-0.0442	-0.0464**	
	(0.0172)	(0.0159)	(0.0195)	(0.0219)	(0.0276)	(0.0218)	
5km							-0.0436**
							(0.0219)
(7) $\ln a_{it}^{\mathrm{MNF_o}}$							
1km	-0.0133	-0.0134	-0.0132	-0.0133	-0.00819	-0.0142	-0.0133
	(0.00886)	(0.00828)	(0.00945)	(0.00922)	(0.00928)	(0.0101)	(0.00922)
(8) $\ln a_{it}^{\mathrm{POP}}$							
20km	1.361	1.429	1.365	1.332	0.966	1.594*	1.332
	(1.039)	(0.896)	(1.030)	(1.050)	(1.078)	(0.895)	(1.050)
I2000	0.820***	0.850***	0.821***	0.814***	0.780***	0.821***	0.814***
	(0.0324)	(0.0283)	(0.0333)	(0.0285)	(0.0390)	(0.0280)	(0.0285)
(10) H. J p-value	0.358	0.348	0.363	0.363	0.352	0.363	0.363
(11) F	266	258.3	249.5	237.7	238	239.6	237.7
(12) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Table E.8: Regression results continued (Dependent variable: $\ln k_{it}^{D}$)

	Novelty (IV3-5)						
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) $\ln \Delta n_{it}$	1.719***	1.734***	1.720***	1.724***	1.717***	1.728***	1.711***
	(0.0840)	(0.0892)	(0.0827)	(0.0786)	(0.0817)	(0.0811)	(0.0761)
(2) $\ln k_{it}$	-0.0339	-0.03 49	-0.0293	-0.0278	-0.0305	-0.0333	-0.0337
	(0.0627)	(0.0624)	(0.0647)	(0.0628)	(0.0642)	(0.0618)	(0.0605)
(3) $(\ln k_{it})^2$	0.261***	0.264***	0.260***	0.261***	0.260***	0.262***	0.259***
	(0.0234)	(0.0224)	(0.0235)	(0.0237)	(0.0234)	(0.0232)	(0.0238)
(4) $\ln a_{it}^{\mathrm{INV}}$							
1km	0.0832	0.0902	0.0800	0.0935	0.0811	0.0756	0.0647
	(0.103)	(0.102)	(0.106)	(0.0974)	(0.102)	(0.114)	(0.108)
(5) $\ln a_{it}^{\mathrm{R\&D}}$							
1km	0.0165*	0.0156*	0.0187*	0.0348**	0.0156	0.0143	0.0125
	(0.00931)	(0.00926)	(0.0107)	(0.0154)	(0.0104)	(0.00888)	(0.00974)
(6) $\ln a_{it}^{\mathrm{MNF_e}}$							
1km			-0.0511	-0.114	-0.0338*	-0.0512*	-0.0425*
			(0.0434)	(0.0779)	(0.0192)	(0.0272)	(0.0221)
10km	-0.107*						
	(0.0572)						
20km		-0.148**					
		(0.0729)					
(7) $\ln a_{it}^{\mathrm{MNF_o}}$							
1km	-0.0147*	-0.00892				-0.0138	-0.0108
	(0.00890)	(0.00783)				(0.00843)	(0.0109)
5km			-0.0346				
			(0.0540)				
10km				-0.219			
				(0.156)			
20km					-0.00425		
					(0.0639)		
(8) $\ln a_{it}^{\mathrm{POP}}$							
5km						0.494	
						(0.529)	
10km							1.445
							(1.530)
20km	1.287	1.390	1.178	-0.112	1.461		
	(1.112)	(1.129)	(1.043)	(1.348)	(1.194)		
(9) τ_1	0.819***	0.834***	0.817***	0.856***	0.807***	0.795***	0.809***
	(0.0296)	(0.0348)	(0.0330)	(0.0489)	(0.0301)	(0.0177)	(0.0318)
(10) H. J p-value	0.361	0.359	0.361	0.359	0.362	0.356	0.348
(11) F	238.6	241.1	238.8	239.2	238.4	242.9	230.5
(12) #Obs.	94,694	94,694	94,694	94,694	94,694	94,694	94,694

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

E.4 IVs based on indirect collaborators in different firms

Tables E.9 and E.10 show the first-stage, and tables E.11 and E.12 show the second-stage regression results for models (4.1) and (4.6), respectively. The choices of IVs in tables E.11 and E.12 are similar to those in tables 7.1 and 7.3, respectively.

Table E.9: First-stage regression results for (4.1) under alternative IVs

	Citations				Novelty			
Variables	(1) IV3-5	(2) IV3	(3) IV4	(4) IV5	(5) IV3-5	(6) IV3	(7) IV4	(8) IV5
(1) k_{it}^{D,IV_3}	0.239***	0.267***			0.251***	0.313***		
	(0.0254)	(0.0110)			(0.0203)	(0.0111)		
(2) k_{it}^{D,IV_4}	0.0592**		0.226***		0.106***		0.314***	
	(0.0235)		(0.0215)		(0.0279)		(0.0150)	
(3) k_{it}^{D,IV_5}	-0.0131			0.163***	0.0282			0.277***
	(0.0493)			(0.0417)	(0.0338)			(0.0267)
(4) $\ln k_{it}$	0.130***	0.130***	0.134***	0.136***	0.158***	0.159***	0.167***	0.170***
	(0.0331)	(0.0330)	(0.0356)	(0.0369)	(0.0126)	(0.0128)	(0.0163)	(0.0163)
(5) $(\ln k_{it})^2$	-0.0513***	-0.0513***	-0.0525***	-0.0532***	-0.0878***	-0.0882***	-0.0891***	-0.0897***
	(0.0117)	(0.0116)	(0.0131)	(0.0136)	(0.00845)	(0.00836)	(0.0102)	(0.0107)
(6) $\ln a_{it}^{\mathrm{INV}}$	0.419***	0.419***	0.428***	0.431***	0.553***	0.561***	0.573***	0.581***
	(0.0777)	(0.0770)	(0.0779)	(0.0784)	(0.104)	(0.105)	(0.108)	(0.112)
(7) $\ln a_{it}^{\mathrm{R\&D}}$	0.00353	0.00387	0.00520	0.00810	0.0145	0.0153	0.0186	0.0242
	(0.0100)	(0.0102)	(0.0105)	(0.0110)	(0.0150)	(0.0157)	(0.0162)	(0.0175)
(8) $\ln a_{it}^{\mathrm{MNF}_e}$	-0.0640**	-0.0640**	-0.0652**	-0.0652**	-0.110***	-0.110***	-0.117***	-0.120***
	(0.0270)	(0.0265)	(0.0295)	(0.0302)	(0.0201)	(0.0177)	(0.0232)	(0.0259)
(9) $\ln a_{it}^{\mathrm{MNF}_o}$	0.0191	0.0198	0.0182	0.0196	0.00403	0.00514	0.00312	0.00233
	(0.0245)	(0.0236)	(0.0255)	(0.0251)	(0.0283)	(0.0260)	(0.0304)	(0.0296)
(10) $\ln a_{it}^{\mathrm{POP}}$	0.652	0.708	0.674	0.855	2.930**	3.063***	2.971**	3.219**
	(1.281)	(1.199)	(1.326)	(1.297)	(1.249)	(1.136)	(1.326)	(1.300)
(11) τ_1	0.318***	0.324***	0.343***	0.369***	0.498***	0.524***	0.549***	0.587***
	(0.0383)	(0.0314)	(0.0396)	(0.0430)	(0.0276)	(0.0300)	(0.0299)	(0.0332)
(12) R^2	0.182	0.181	0.172	0.165	0.194	0.193	0.185	0.177
(13) F	268.1	593.7	110	15.33	494	799	437.2	108.1
(14) p-value	0	0	0	0.000129	0	0	0	0
(15) #Obs.	103,862	103,862	103,862	103,862	103,862	103,862	103,862	103,862

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) inventor, IPC class and period fixed effects are controlled. (*iii*) ***p<0.01, **p<0.05, * p<0.1

Table E.10: First-stage regression results for (4.6) under alternatie IVs

	Citations				Novelty			
Variables	(1) IV3-5	(2) IV3	(3) IV4	(4) IV5	(5) IV3-5	(6) IV3	(7) IV4	(8) IV5
(1) $\ln \Delta n_{it}^{\mathrm{IV}_3}$	0.247***	0.233***			0.247***	0.233***		
	(0.00775)	(0.00780)			(0.00775)	(0.00780)		
(2) $\ln \Delta n_{it}^{\mathrm{IV}_4}$	-0.0345**		0.219***		-0.0345**		0.219***	
	(0.0153)		(0.00834)		(0.0153)		(0.00834)	
(3) $\ln \Delta n_{it}^{\mathrm{IV}_5}$	0.0209			0.200***	0.0209			0.200***
	(0.0160)			(0.00837)	(0.0160)			(0.00837)
(4) $\ln k_{it}$	0.0655**	0.0657**	0.0746***	0.0811***	0.0655**	0.0657**	0.0746***	0.0811***
	(0.0268)	(0.0267)	(0.0285)	(0.0296)	(0.0268)	(0.0267)	(0.0285)	(0.0296)
(5) $(\ln k_{it})^2$	-0.165***	-0.165***	-0.171***	-0.175***	-0.165***	-0.165***	-0.171***	-0.175***
	(0.00833)	(0.00829)	(0.00939)	(0.00999)	(0.00833)	(0.00829)	(0.00939)	(0.00999)
(6) $\ln a_{it}^{\mathrm{INV}}$	0.136***	0.135***	0.134***	0.156**	0.136***	0.135***	0.134***	0.156**
	(0.0433)	(0.0430)	(0.0511)	(0.0615)	(0.0433)	(0.0430)	(0.0511)	(0.0615)
(7) $\ln a_{it}^{\mathrm{R\&D}}$	0.00555	0.00583	0.00413	0.000824	0.00555	0.00583	0.00413	0.000824
	(0.00561)	(0.00546)	(0.00584)	(0.00599)	(0.00561)	(0.00546)	(0.00584)	(0.00599)
(8) $\ln a_{it}^{\mathrm{MNF}_e}$	0.0210*	0.0217*	0.0233	0.0126	0.0210*	0.0217*	0.0233	0.0126
	(0.0127)	(0.0123)	(0.0144)	(0.0189)	(0.0127)	(0.0123)	(0.0144)	(0.0189)
(9) $\ln a_{it}^{\mathrm{MNF}_o}$	0.0106	0.0104	0.0133	0.0164	0.0106	0.0104	0.0133	0.0164
	(0.00826)	(0.00797)	(0.00856)	(0.0104)	(0.00826)	(0.00797)	(0.00856)	(0.0104)
(10) $\ln a_{it}^{\mathrm{POP}}$	0.403	0.409	0.568**	0.671**	0.403	0.409	0.568**	0.671**
	(0.278)	(0.280)	(0.281)	(0.304)	(0.278)	(0.280)	(0.281)	(0.304)
(11) τ_1	-0.177***	-0.175***	-0.189***	-0.202***	-0.177***	-0.175***	-0.189***	-0.202***
	(0.0140)	(0.0132)	(0.0128)	(0.0142)	(0.0140)	(0.0132)	(0.0128)	(0.0142)
(12) R^2	0.329	0.329	0.305	0.284	0.329	0.329	0.305	0.284
(13) F	470	892	692	571	470	892	692	571
(14) p-value	0	0	0	0	0	0	0	0
(15) #Obs.	88,204	88,204	88,204	88,204	88,204	88,204	88,204	88,204

(*i*) Standarderrors clustered by UAs are in parentheses. (*ii*) inventor, IPC class and period fixed effects are controlled. (*iii*) ***p<0.01, **p<0.05, * p<0.1

Table E.11: Regression results for (4.1) under alternative IVs

	Citations					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln k^{D}_{it}$	0.166***	0.225***	0.240***	0.0980	-0.0272	0.167***	0.284***	0.288***	0.268***	0.291***
	(0.0106)	(0.0455)	(0.0404)	(0.111)	(0.191)	(0.00529)	(0.0412)	(0.0418)	(0.0444)	(0.0702)
(2) $\ln k_{it}$	0.115***	0.107***	0.105***	0.125***	0.142***	0.155***	0.134***	0.133***	0.137***	0.133***
	(0.0165)	(0.0198)	(0.0183)	(0.0339)	(0.0497)	(0.0169)	(0.0237)	(0.0236)	(0.0242)	(0.0270)
(3) $(\ln k_{it})^2$	-0.0887***	-0.0855***	-0.0847***	-0.0924***	-0.0992***	-0.194***	-0.183***	-0.183***	-0.185***	-0.183***
	(0.0105)	(0.0123)	(0.0118)	(0.0174)	(0.0231)	(0.00999)	(0.00675)	(0.00641)	(0.00742)	(0.0119)
(4) $\ln a^{\mathrm{INV}}_{it}$	0.187***	0.161***	0.154**	0.217***	0.272***	0.334***	0.262***	0.259***	0.272***	0.258***
	(0.0603)	(0.0613)	(0.0618)	(0.0684)	(0.0919)	(0.0927)	(0.0940)	(0.0966)	(0.0912)	(0.0582)
(5) $\ln a^{\mathrm{R\&D}}_{it}$	0.0279***	0.0274***	0.0273***	0.0285***	0.0295***	0.0434***	0.0403***	0.0402***	0.0407***	0.0401***
	(0.00781)	(0.00734)	(0.00723)	(0.00838)	(0.00945)	(0.0151)	(0.0132)	(0.0132)	(0.0133)	(0.0121)
(6) $\ln a^{\mathrm{MNF}e}_{it}$	0.0117*	0.0156***	0.0166***	0.00724	-0.00102	-0.0130	0.00138	0.00189	-0.000597	0.00222
	(0.00695)	(0.00576)	(0.00559)	(0.00774)	(0.0100)	(0.0115)	(0.00923)	(0.00884)	(0.00978)	(0.0164)
(7) $\ln a^{\mathrm{MNF}o}_{it}$	0.00630	0.00492	0.00457	0.00788	0.0108*	-0.00250	-0.00317	-0.00319	-0.00308	-0.00321
	(0.00670)	(0.00716)	(0.00755)	(0.00481)	(0.00616)	(0.00642)	(0.00801)	(0.00808)	(0.00778)	(0.00804)
(8) $\ln a^{\mathrm{POP}}_{it}$	-0.644	-0.713	-0.731	-0.565	-0.419	0.818	0.405	0.390	0.462	0.381
	(0.533)	(0.514)	(0.508)	(0.601)	(0.731)	(0.503)	(0.495)	(0.504)	(0.488)	(0.411)
(9) τ_1	0.225***	0.201***	0.195***	0.253***	0.306***	0.303***	0.222***	0.219***	0.233***	0.217***
	(0.0168)	(0.0257)	(0.0231)	(0.0562)	(0.0934)	(0.0346)	(0.0545)	(0.0550)	(0.0557)	(0.0610)
(10) R^2	0.153					0.186				
(11) Hansen J p-val.		0.494					0.373			
(12) 1st stage F		322.6	944.7	522.2	209		372.1	1041	694.2	387.1
(13) #Obs.	103,862	103,862	103,862	103,862	103,862	103,862	103,862	103,862	103,862	103,862

(*i*) standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.

Table E.12: Regression results for (4.6) under alternative IVs

	Citations					Novelty				
Variables	(1) OLS	(2) IV3-5	(3) IV3	(4) IV4	(5) IV5	(6) OLS	(7) IV3-5	(8) IV3	(9) IV4	(10) IV5
(1) $\ln \Delta n_{it}$	0.0993***	0.900***	0.901***	0.956***	0.986***	0.236***	1.216***	1.217***	1.280***	1.316***
	(0.00687)	(0.0293)	(0.0299)	(0.0352)	(0.0301)	(0.00901)	(0.0482)	(0.0491)	(0.0621)	(0.0611)
(2) $\ln k_{it}$	0.133***	0.0365	0.0363	0.0296	0.0260	0.152***	0.0341	0.0339	0.0263	0.0221
	(0.0423)	(0.0562)	(0.0562)	(0.0567)	(0.0580)	(0.0328)	(0.0483)	(0.0482)	(0.0480)	(0.0492)
(3) $(\ln k_{it})^2$	-0.0390***	0.123***	0.123***	0.135***	0.141***	-0.0458***	0.153***	0.153***	0.166***	0.173***
	(0.0148)	(0.0162)	(0.0160)	(0.0153)	(0.0165)	(0.0142)	(0.0135)	(0.0133)	(0.0114)	(0.0127)
(4) $\ln a^{\mathrm{INV}}_{it}$	0.376***	0.146***	0.146***	0.130***	0.121***	0.501***	0.220***	0.219***	0.201**	0.191**
	(0.0928)	(0.0354)	(0.0352)	(0.0324)	(0.0319)	(0.114)	(0.0777)	(0.0779)	(0.0794)	(0.0778)
(5) $\ln a^{\mathrm{R\&D}}_{it}$	0.0100	0.00329	0.00328	0.00281	0.00256	0.0290*	0.0208**	0.0208**	0.0202**	0.0199**
	(0.0109)	(0.00561)	(0.00560)	(0.00541)	(0.00528)	(0.0165)	(0.0102)	(0.0102)	(0.01000)	(0.00984)
(6) $\ln a^{\mathrm{MNF}e}_{it}$	-0.0680**	-0.0322**	-0.0321**	-0.0296**	-0.0283*	-0.109***	-0.0653***	-0.0653***	-0.0624***	-0.0608***
	(0.0269)	(0.0151)	(0.0151)	(0.0149)	(0.0146)	(0.0219)	(0.0150)	(0.0150)	(0.0154)	(0.0159)
(7) $\ln a^{\mathrm{MNF}o}_{it}$	0.0210	0.0118	0.0118	0.0112	0.0108	0.000656	-0.0106	-0.0106	-0.0113	-0.0118
	(0.0231)	(0.0115)	(0.0115)	(0.0109)	(0.0107)	(0.0266)	(0.0106)	(0.0106)	(0.00974)	(0.00940)
(8) $\ln a^{\mathrm{POP}}_{it}$	1.018	-0.161	-0.163	-0.245	-0.289	3.127***	1.683*	1.681*	1.588*	1.536
	(1.124)	(1.130)	(1.129)	(1.136)	(1.148)	(1.194)	(0.960)	(0.959)	(0.945)	(0.950)
(9) τ_1	0.393***	0.456***	0.456***	0.460***	0.463***	0.668***	0.745***	0.746***	0.751***	0.753***
	(0.0325)	(0.0443)	(0.0443)	(0.0455)	(0.0457)	(0.0312)	(0.0269)	(0.0269)	(0.0279)	(0.0280)
(10) R^2	0.159					0.175				
(11) Hansen J p-val.		0.194					0.185			
(12) 1st stage F		2800	8401	6651	5010		2800	8401	6651	5010
(13) #Obs.	88,204	88,204	88,204	88,204	88,204	88,204	88,204	88,204	88,204	88,204

(*i*) Standard errors clustered by UAs are in parentheses. (*ii*) IPC class fixed effects are controlled. (*iii*) *** p<0.01, ** p<0.05, * p<0.1.